OPIC
Office de la propriété intellectuelle du Canada

CIPO
Canadian Intellectual Property Office

(12) (19) (CA) **Demande-Application**

(21) (A1) **2,251,262**
(22) 1998/11/18
(43) 1999/05/20

(72) DIETRICH, Paul Shartzer, US
(72) FISH, Linda Marie, US
(72) KHARE, Reena, US
(72) RABERT, Douglas Kenneth, US
(72) SANGAMESWARAN, Lakshmi, US
(71) F.HOFFMANN-LA ROCHE AG, CH

(51) Int.Cl.[6] C12N 15/12, C07K 14/705, C12Q 1/68, G01N 33/68, C07K 16/28
(30) 1997/11/20 (60/066,225) US

(54) **CODAGE D'ACIDE NUCLEIQUE POUR CANAL SODIQUE DE TISSU NERVEUX**

(54) **NUCLEIC ACID ENCODING A NERVOUS TISSUE SODIUM CHANNEL**

(57) A novel nucleic acid sequence encoding for a mammalian voltage-gated, preferably TTX-resistant, sodium channel is isolated. Also disclosed are polypeptide products of recombinant expression of these sequences, expression vectors comprising the DNA sequence, and host cells transformed with these expression vectors. Other aspects of this invention are peptides whose sequences are based on the amino acid sequences deduced from these DNA sequences, antibodies specific for such proteins and peptides, procedures for detection and quantitation of such proteins and nucleic acids related thereto. Another aspect of this invention is the use of this voltage-gated, preferably tetrodotoxin-resistant, sodium channel as a therapeutic target for compounds.

ABSTRACT OF THE DISCLOSURE

A novel nucleic acid sequence encoding for a mammalian voltage-gated, preferably TTX-resistant, sodium channel is isolated. Also disclosed are polypeptide products of recombinant expression of these sequences, expression vectors comprising the DNA sequence, and host cells transformed with these expression vectors. Other aspects of this invention are peptides whose sequences are based on the amino acid sequences deduced from these DNA sequences, antibodies specific for such proteins and peptides, procedures for detection and quantitation of such proteins and nucleic acids related thereto. Another aspect of this invention is the use of this voltage-gated, preferably tetrodotoxin-resistant, sodium channel as a therapeutic target for compounds.

This invention relates generally to sodium channel proteins and more particularly to a novel nucleic acid sequence encoding for a mammalian α-subunit of a voltage-gated, preferably tetrodotoxin-resistant, nervous tissue sodium channel protein. This invention further relates to its production by recombinant technology.

The basic unit of information transmitted from one part of the nervous system to another is a single action potential or nerve impulse. The „transmission line" for these impulses is the axon, or nerve fiber. The electrical excitability of the nerve membrane has been shown to depend on the membrane's voltage-sensitive ionic permeability system that allows it to use energy stored in ionic concentration gradients. Electrical activity of the nerve is triggered by a depolarization of the membrane, which opens channels through the membrane that are highly selective for sodium ions, which are then driven inward by the electrochemical gradient. Of the many ionic channels, the voltage-gated or voltage-sensitive sodium channel is one of the most studied. It is a transmembrane protein that is essential for the generation of action potentials in excitable cells. An excellent review of sodium channels is presented in Catterall, TINS 16(12), 500-506 (1993).

The cDNAs for several Na^+ channels have been cloned and sequenced. Numa *et al.*, Annals of the New York Academy of Sciences 479, 338-355 (1986), describe cDNA from the electric organ of eel and two different ones from rat brain. Rogart, U.S. Patent No. 5,380,836, describes cDNA from rat cardiac tissue. See also Rogart *et al.*, Proc. Natl. Acad. Sci. 86, 8170-8174 (1989). The sequence of PN1 and its orthologs in humans (hNE) and rabbits (Na^+s) have been published (see, for example, Klugbauer *et al.*, EMBOJ 14, 1084-1090 (1995) and Belcher *et al.*, Proc. Natl. Acad. Sci. U.S.A. 923, 11034-11038 (1995)). The sequence of rat PN1 cloned from DRG and its function expression have been described (see, for example, Sangameswaran *et al.*, J.Biol.Chem. 272, 14805-14809 (1997)). Other cloned sodium channels include rat brain types I and II, Noda *et al.*, Nature 320, 188-192 (1986), IIa, Auld *et al.*, Neuron 1, 449-461 (1988), and III, Kayano *et al.*, FEBS Lett. 228, 187-194 (1988), rat

skeletal muscle (SkM1), Trimmer *et al.*, Neuron 3, 33-49 (1989), rat NaCh6, Schaller *et al.*, J. Neurosci. 15, 3231-3242 (1995), rat peripheral nerve sodium channel type 3 (rPN3), Sangameswaran *et al.*, J. Biol Chem. 271, 5953-5956 (1996), also called SNS, Akopian *et al.*, Nature 379, 257-262 (1996), rat atypical channel, Felipe *et al.*, J. Biol. Chem. 269, 30125-30131 (1994), and the rat glial sodium channel, Akopian *et al.*, FEBS Lett. 400, 183-187 (1997).

These studies have shown that the amino acid sequence of the Na^+ channel has been conserved over a long evolutionary period. These studies have also revealed that the channel is a single polypeptide containing four internal repeats, or homologous domains (domains I-IV), having similar amino acid sequences. Each domain folds into six predicted and helical transmembrane segments: five are hydrophobic segments and one is highly charged with many positively charged lysine and arginine residues. This highly charged segment is the fourth transmembrane segment in each domain (the S4 segment) and is likely to be involved in voltage-gating. The positively charged side chains on the S4 segment are likely to be paired with the negatively charged side chains on the other five segments such that membrane depolarization could shift the position of one helix relative to the other, thereby opening the channel. Accessory subunits may modify the function of the channel.

Therapeutic utility in recombinant materials derived from the DNA of the numerous sodium channels have been discovered. For example, U.S. Patent No. 5,132,296 by Cherksey discloses purified Na^+ channels that have proven useful as therapeutic and diagnostic tools.

Isoforms of sodium channels are divided into „subfamilies". The term „isoform" is used to mean distinct but closely related sodium channel proteins, i.e., those having an amino acid homology of approximately 60-80%. These also show strong homology in functions. The term „subfamilies" is used to mean distinct sodium channels that have an amino acid homology of approximately 80-95%. Combinations of several factors are used to determine the distinctions within a subfamily, for example, the speed of a channel, chromosomal location, expression data, homology to other channels within a species, and homology to a

channel of the same subfamily across species. Another consideration is an affinity to tetrodotoxin („TTX"). TTX is a highly potent toxin from the puffer or fugu fish which blocks the conduction of nerve impulses along axons and in excitable membranes of nerve fibers. TTX binds to the Na^+ channel and blocks the flow of sodium ions.

Studies employing TTX as a probe have shed much light on the mechanism and structure of Na^+ channels. There are three Na^+ channel subtypes that are defined by the affinity for TTX, which can be measured by the IC_{50} values: TTX-sensitive Na^+ channels (IC_{50} ≈ 1-30 nM), TTX-insensitive Na^+ channels (IC_{50} ≈ 1-5 μM), and TTX-resistant Na^+ channels (IC_{50} ≥ 50 μM).

TTX-insensitive action potentials were first studied in rat skeletal muscle (Redfern *et al.*, Acta Physiol. Scand. 82, 70-78 (1971)). Subsequently, these action potentials were described in other mammalian tissues, including newborn mammalian skeletal muscle, mammalian cardiac muscle, mouse dorsal root ganglion cells in vitro and in culture, cultured mammalian skeletal muscle and L6 cells. See Rogart, Ann. Rev. Physiol. 43, 711-725 (1980).

Rat dorsal root ganglia neurons possess both TTX-sensitive (IC_{50} ~ 0.3 nM) and TTX-resistant (IC_{50} ~ 100 μM) sodium channel currents, as described in Roy *et al.*, J. Neurosci. 12, 2104-2111 (1992). TTX-resistant sodium currents have also been measured in rat nodose and petrosal ganglia. See Ikeda *et al.*, J. Neurophysiol. 55, 527-539 (1986) and Stea *et al.*, Neurosci. 47, 727-736 (1992). Electrophysiologists believe that another TTX-resistant sodium channel is yet to be detected.

Though cDNAs from rat skeletal muscle, heart and brain are known, identification and isolation of cDNA from peripheral sensory nerve tissue, such as dorsal root ganglia, has been hampered by the difficulty of working with such tissue.

SUMMARY OF THE INVENTION

The present invention provides novel purified and isolated nucleic acid sequences encoding mammalian, preferably TTX-resistant, nervous tissue sodium channel proteins that

are strongly expressed in adult DRG and nodose ganglia, less strongly expressed in brain, spinal cord and superior cervical ganglia, and not expressed in sciatic nerve, heart or skeletal muscle. In presently preferred forms, novel DNA sequences comprise cDNA sequences encoding rat nervous tissue sodium channel protein. One aspect of the present invention is the α-subunit of this sodium channel protein.

Disclosed is the DNA, cDNA, and mRNA derived from the nucleic acid sequences of the invention and the cRNA derived from the mRNA. Specifically, two cDNA sequences together encode for the full length rat nervous tissue sodium channel.

Also included in this invention are alternate DNA forms, such as genomic DNA, DNA prepared by partial or total chemical synthesis from nucleotides, and DNA having deletions or mutations.

Still another aspect of the invention is the novel rat TTX-resistant sodium channel protein and fragments thereof, encoded by the DNA of this invention.

Another aspect of the present invention are recombinant polynucleotides and oligonucleotides comprising a nucleic acid sequence derived from the DNA sequence of this invention.

Another aspect of the invention is a method of stabilizing the full length cDNA which encodes the protein sequence of the invention.

Further aspects of the invention include expression vectors comprising the DNA of the invention, host cells transformed or transfected by these vectors, and a cDNA library of these host cells.

Also forming part of this invention is an assay for inhibitors of the sodium channel protein comprising contacting a compound suspected of being an inhibitor with expressed sodium channel and measuring the activity of the sodium channel.

Further provided is a method of inhibiting the activity of the TTX-resistant sodium channel comprising administering an effective amount of a compound having an IC_{50} of 10 μM or less.

Additionally provided are methods of employing the DNA for forming monoclonal and polyclonal antibodies, for use as molecular targets for drug discovery, highly specific markers for specific antigens, detector molecules, diagnostic assays, and therapeutic uses, such as pain relief, a probe for the PN5 channel in other mammalian tissue, designing therapeutics and screening for therapies.

BRIEF DESCRIPTION OF THE SEQ ID'S AND FIGURES

Figures 1A-E depict the 5908 nucleotide cDNA native sequence encoding the rat sodium channel type 5 („PN5") (SEQ ID NO: 1), derived from two overlapping cDNA clones, designated 26.2 and 1.18.

Figures 2A-F depict the deduced amino acid sequence of PN5 (SEQ ID NO: 2, represented in the three-letter amino acid code). Figures 2G-H, depicting the deduced amino acid sequence of PN5 in single letter amino acid code, also show the homologous domains (I-IV); the putative transmembrane segments (Sl-S6); the amino acid conferring resistance to TTX (♦); N-glycosylation sites (•); cAMP-dependent protein kinase A (PKA) phosphorylation site (0); and the termination codon (*).

Figure 3A depicts an 856 base pair sequence for the human PN5 (SEQ ID NO: 3). Figure 3B depicts the amino acid sequence comparison of the hPN5 fragment with rat PN5.

Figure 4 depicts the sequence for the novel sodium channel domain IV probe (SEQ ID NO: 4).

Figures 5A-E depict the 5334 nucleotide sequence modified for stability and expression (SEQ ID NO: 5). Nucleotides 24 to 5518 constitute the 5295 bp region coding for a 1765 amino acid protein.

Figure 6 depicts the cloning map of PN5.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a purified and isolated nucleic acid sequence encoding for a novel mammalian, preferably TTX-resistant, sodium channel protein. The term "purified

and isolated DNA" refers to DNA that is essentially free, i.e. contains less than about 30%, preferably less than about 10%, and even more preferably less than about 1%, of the DNA with which the DNA of interest is naturally associated. Techniques for assessing purity are well known to the art and include, for example, restriction mapping, agarose gel electrophoresis, and CsCl gradient centrifugation.

The term "DNA" is meant to include „cDNA", or complementary DNA, which is single-stranded or double-stranded DNA sequences made by reverse transcription of mRNA isolated from a donor cell or by chemical synthesis. For example, treatment of mRNA with a reverse transcriptase such as AMV reverse transcriptase or M-MuLV reverse transcriptase in the presence of an oligonucleotide primer will furnish an RNA-DNA duplex which can be treated with RNase H, DNA polymerase, and DNA ligase to generate double-stranded cDNA. If desired, the double-stranded cDNA can be denatured by conventional techniques such as heating to generate single-stranded cDNA. The term „cDNA" includes cDNA that is a complementary copy of the naturally occurring mRNA ,as well as complementary copies of variants of the naturally occurring mRNA that have the same biological activity. Variants would include, for example, insertions, deletions, sequences with degenerate codons and alleles.

„cRNA" corresponding to mRNA transcribed from a DNA sequence encoding the α-subunit of a novel, preferably TTX-resistant, sodium channel protein is contemplated by this invention. The term „cRNA" refers to RNA that is a copy of the mRNA transcribed by a cell.

Specifically, the invention encompasses DNA having the native versions of the nucleotide sequences set forth in Figures 1A-E (SEQ ID NO: 1) designated herein as sodium channel type 5 (PN5). Figures 1A-E depict the 5908 nucleotide cDNA construct comprising a 5298-base (counting the stop codon) open reading frame (SEQ ID NO:1). Nucleotide residue 79 represents the start site of translation and residue 5376 represents the end of the stop codon.

The invention also encompasses engineered versions of PN5, and specifically the version as set forth in Figures 5A-E (SEQ ID NO: 5). This 5334 nucleotide SaII-XbaI clone

lacks most of the untranslated sequences, the 5298 nucleotide open reading frame beginning at nucleotide 24 and ending at nucleotide 5321. The start and stop codons are underlined, as are the translationally silent mutations at nucleotides 3932, 3935, 3941, 3944, and 3947, which were introduced to block rearrangement in this region during growth in *E. Coli*.

The nucleotide sequence of SEQ ID NO: 1 (Figures 1A-E) corresponds to the cDNAs from rat. A homology search provided that the closest related sodium channel is found in the rat cardiac channel, with 72.5% homology. The next closely related channels are rPN1, with 72% and rat brain types I and III, with 71.8% and 71.3% respectively. Homology to rPN3a, hPN3, rPN4, rPN4a, rat brain type II and rat skeletal muscle are each approximately 70 to 71%.

Additionally, an 856 base pair clone (SEQ ID NO: 3) as shown in Figure 3A has been isolated from a human dorsal root ganglia (DRG) „cDNA library" and is closely related to the rat PN5 amino acid sequence with 79% identity and 86% homology. The human PN5 sequence spans the region between IIIS1 and interdomain III/IV which includes the fast inactivation gate (i.e., IFM) that is located within interdomain III/IV.

The term „cDNA library" refers to a collection of clones, usually in a bacteriophage, or less commonly in bacterial plasmids, containing cDNA copies of mRNA sequences derived from a donor cell or tissue.

It is believed that additional homologs of the novel rat TTX-resistant sodium channel described herein are also expressed in other mammalian tissue.

Northern blot analysis (Example 5) indicates that PN5 is encoded by a ~6.5 kb transcript.

The deduced amino acid sequence of PN5, shown in Figures 2A-F (SEQ ID NO: 2), exhibits the primary structural features of an α-subunit of a voltage-gated, TTX-resistant sodium channel. Shown in Figures 2G-H are the homologous domains (I-IV); the putative transmembrane segments (Sl-S6); the amino acid conferring resistance to TTX (♦); N-glycosylation sites (•); and cAMP-dependent PKA phosphorylation sites (0). DNA sequences

encoding the same or allelic variant or analog sodium channel protein polypeptides of the nervous system, through use of, at least in part, degenerate codons are also contemplated by this invention.

An interesting feature of this deduced amino acid sequence is that the amino acid that is most responsible for TTX-sensitivity is located at position 355 and is not aromatic. In rat and human brain type sodium channels, skeletal muscle channel, and in PN1 and PN4, this amino acid is tyrosine or phenylalanine and these channels are all TTX-sensitive. In PN3 and PN5, the amino acid is a serine. Since PN3 is highly resistant to TTX, the implication is that PN5 is also a TTX-resistant channel. The cardiac channel has a cysteine at this position and is „insensitive" to TTX.

Although PN5 contains all of the hallmark features of a voltage-gated sodium channel, it has unique structural features that distinguish it from other sodium channels. For example, DIIS4 has 5 basic amino acids conserved in all sodium channels that could play a significant role in the voltage sensing aspects of the channel function. In PN5, the first basic amino acid is replaced by an alanine. Similarly, in DIIIS4, PN5 has 5 basic amino acids rather than six that are present in other sodium channel sequences, the last arginine replaced by a glutamine. In DIIIS3, the transmembrane segment contains only 18 amino acids, in contrast to 22 amino acids in other channels. Also, the short linker (4 amino acids) loop between S3 and S4 in DIII is even shorter by a ‚deletion' of 3 amino acids. This shortening of the S3 and the linker loop has been confirmed by designing primers in the appropriate region of the sequence for an RT-PCR experiment from rat DRG and sequencing the amplified DNA fragment. Such an experiment has been performed to confirm the sequence of another region of PN5, in the DIVS5-S6 loop, where there was a deletion of an 8 amino acid peptide.

Reverse transcription-polymerase chain reaction (oligonucleotide-primed RT-PCR) tissue distribution analysis of RNA from the rat central and peripheral nervous systems, in particular from rat DRG, was performed. Eight main tissue types were screened for expression of the unique PN5 genes corresponding to positions 5651-5903 of SEQ ID NO: 1

(Figures 1A-E). PN5 mRNA was present in five of the tissues studied: brain, spinal cord, DRG, nodose ganglia, and superior cervical ganglia. PN5 was not present in the remaining tissues studied: sciatic nerve tissue, heart or skeletal muscle tissue. PN5 was found to be the strongest in DRG and nodose ganglia, leading the applicants to believe that the DRG is enriched with PN5. PN5 shows dramatic abundance differences across a range of tissues. PN5 has a gradient of expression with high expression in DRG. PN5 has a gradient of expression like other channels, but more limited distribution.

The invention not only includes the entire protein expressed by the cDNA sequences of SEQ ID NOS: 1, 2 and 3, but also includes protein fragments. These fragments can be obtained by cleaving the full length proteins or by using smaller DNA sequences or „polynucleotides“ to express the desired fragment.

The term "polynucleotide“ as used herein refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA, as well as double- and single-stranded RNA. It also includes modified, for example, by methylation and/or by capping, and unmodified forms of the polynucleotide.

Further, the term "polynucleotide“ is intended to include a recombinant polynucleotide, which is of genomic, cDNA, semisynthetic or synthetic origin which, by virtue of its origin or manipulation is not associated with all or a portion of the polynucleotide with which it is associated in nature and/or is linked to a polynucleotide other than that to which it is linked in nature.

Accordingly, the invention also includes polynucleotides that can be used to make polypeptides of about 10 to 1500, preferably 10 to 100, amino acids in length. The isolation and purification of such recombinant polypeptides can be accomplished by techniques that are well known in the art, for example, preparative chromatographic separations or affinity chromatography. In addition, polypeptides can also be made by synthetic means which are well known in the art.

The invention allows for the manipulation of genetic materials by recombinant technology to produce polypeptides that possess the structural and functional characteristics of the novel voltage-gated, TTX-resistant sodium channel α-subunit found in sensory nerves. Site directed mutagenesis can be used to provide such recombinant polypeptides. For example, synthetic oligonucleotides can be specifically inserted or substituted into the portion of the gene of interest to produce genes encoding for and expressing a specific mutant. Random degenerate oligonucleotides can also be inserted and phage display techniques can be used to identify and isolate polypeptides possessing a functional property of interest.

In addition, the present invention contemplates recombinant polynucleotides of about 15 to 20kb, preferably 10 to 15kb, nucleotides in length, comprising a nucleic acid sequence „derived from" the DNA of the invention.

The term "derived from" a designated sequence, refers to a nucleic acid sequence that is comprised of a sequence of approximately at least 6 to 8 nucleotides, more preferably at least 10 to 12 nucleotides, and, even more preferably, at least 15 to 20 nucleotides that correspond to, i.e., are homologous or complementary to, a region of the designated sequence. The derived sequence is not necessarily physically derived from the nucleotide sequence shown, but may be derived in any manner, including for example, chemical synthesis or DNA replication or reverse transcription, which are based on the information provided by the sequences of bases in the region(s) from which the polynucleotide is derived.

A neonatal expression test was performed with F11, a fusion cell line designed from neonatal rat DRG fused with a mouse cell line, N18TG, from Massachusetts General Hospital. F11 responds to trophic agents, such as NGF, by extending dendrites. It was found that PN5 was present in both native F11 and F11 treated with NGF, leading the applicants to believe that the sodium channel is natively expressed in F11.

In situ hybridization of PN5 mRNA to rat DRG tissue provides localization predominantly in the small and medium neurons with no detection in large neurons.

PN5 was also mapped to its cytogenetic location on mouse chromosome preparations. PN5 maps to the same chromosome as the cardiac channel and PN3.

In general, sodium channels comprise an α- and two ß-subunits. The ß-subunits may modulate the function of the channel. However, since the α-subunit is all that is required for the channel to be fully functional, expression of the cDNA in SEQ ID NO: 1 (Figures 1A-E) will provide a fully functional protein. The gene encoding the $ß_1$-subunit in peripheral nerve tissue was found to be identical to that found in rat heart, brain and skeletal muscle. The cDNA of the $ß_1$-subunit is not described herein as it is well known in the art, see Isom *et al.*, Neuron 12, 1183-1194 (1994). However, it is to be understood that by combining the known sequence for the $ß_1$-subunit with the α-subunit sequence described herein, one may obtain complete PN5 voltage-gated, preferably TTX-resistant, sodium channel.

The present invention also includes „expression vectors" comprising the DNA or the cDNA described above, host cells transformed with these expression vectors capable of producing the sodium channel of the invention, and cDNA libraries comprising such host cells.

The term "expression vector" refers to any genetic element, e.g., a plasmid, a chromosome, a virus, behaving either as an autonomous unit of polynucleotide expression within a cell or being rendered capable of replication by insertion into a host cell chromosome, having attached to it another polynucleotide segment, so as to bring about the replication and/or expression of the attached segment. Suitable vectors include, but are not limited to, plasmids, bacteriophages, and cosmids. Vectors will contain polynucleotide sequences which are necessary to effect ligation or insertion of the vector into a desired host cell and to effect the expression of the attached segment. Such sequences differ depending on the host organism, and will include promoter sequences to effect transcription, enhancer sequences to increase transcription, ribosomal binding site sequences and transcription and translation termination sequences.

The term "host cell" generally refers to prokaryotic or eukaryotic organisms and includes any transformable or transfectable organism which is capable of expressing a protein and can be, or has been, used as a recipient for expression vectors or other transferred DNA. Host cells can also be made to express protein by direct injection with exogenous cRNA translatable into the protein of interest. A preferred host cell is the *Xenopus* oocyte.

The term "transformed" refers to any known method for the insertion of foreign DNA or RNA sequences into a host prokaryotic cell. The term „transfected" refers to any known method for the insertion of foreign DNA or RNA sequences into a host eukaryotic cell. Such transformed or transfected cells include stably transformed or transfected cells in which the inserted DNA is rendered capable of replication in the host cell. They also include transiently expressing cells which express the inserted DNA or RNA for limited periods of time. The transformation or transfection procedure depends on the host cell being transformed. It can include packaging the polynucleotide in a virus as well as direct uptake of the polynucleotide, such as, for example, lipofection or microinjection. Transformation and transfection can result in incorporation of the inserted DNA into the genome of the host cell or the maintenance of the inserted DNA within the host cell in plasmid form. Methods of transformation are well known in the art and include, but are not limited to, viral infection, electroporation, lipofection, and calcium phosphate mediated direct uptake.

It is to be understood that this invention is intended to include other forms of expression vectors, host cells, and transformation techniques which serve equivalent functions and which become known to the art hereto.

The invention also pertains to an assay for inhibitors of the novel TTX-resistant sodium channel protein comprising contacting a compound suspected of being an inhibitor with expressed sodium channel and measuring the activity of the sodium channel. The compound can be a substantially pure compound of synthetic origin combined in an aqueous medium, or the compound can be a naturally occurring material such that the assay medium is an extract of biological origin, such as, for example, a plant, animal, or microbial cell extract.

PN5 activity can be measured by methods such as electrophysiology (two electrode voltage clamp or single electrode whole cell patch clamp), guanidinium ion flux assays, and toxin-binding assays. An "inhibitor" is defined as generally that amount that results in greater than 50% decrease in PN5 activity, preferably greater than 70% decrease in PN5 activity, more preferably greater than 90% decrease in PN5 activity.

Many uses of the invention exist, a few of which are described below:

1. Probe for mamalian channels.

As mentioned above, it is believed that additional homologs of the novel rat TTX-resistant sodium channel described herein are also expressed in mammalian tissue, in particular, human tissue. The entire cDNAs of PN5 rat sodium channels of the present invention can be used as a probe to discover whether additional novel PN5 voltage-gated, preferably TTX-resistant, sodium channels exist in human tissue and, if they do, to aid in isolating the cDNAs for the human protein.

The human homologues of the rat TTX-resistant PN5 channels can be cloned using a human DRG cDNA library. Human DRG are obtained at autopsy. The frozen tissue is homogenized and the RNA extracted with guanidine isothiocyanate (Chirgwin *et al.* Biochemistry 18, 5294-5299, (1979)). The RNA is size-fractionated on a sucrose gradient to enrich for large mRNAs because the sodium channel α-subunits are encoded by large (7-11 kb) transcripts. Double-stranded cDNA is prepared using the SuperScript Choice cDNA kit (GIBCO BRL) with either oligo(dT) or random hexamer primers. EcoRI adapters are ligated onto the double-stranded cDNA which is then phosphorylated. The cDNA library is constructed by ligating the double-stranded cDNA into the bacteriophage-lambda ZAP II vector (Stratagene) followed by packaging into phage particles.

Phage are plated out on 150 mm plates on a lawn of XLI-Blue MRF′ bacteria (Stratagene) and plaque replicas are made on Hybond N nylon membranes (Amersham). Filters are hybridized to rat PN5 cDNA probes by standard procedures and detected by autoradiography or chemiluminescence. The signal produced by the rat PN5 probes

hybridizing to positive human clones at high stringency should be stronger than obtained with rat brain sodium channel probes hybridizing to these clones. Positive plaques are further purified by limiting dilution and re-screened by hybridization or PCR. Restriction mapping and polymerase chain reaction will identify overlapping clones that can be assembled by standard techniques into the full-length human homologue of rat PN5. The human clone can be expressed by injecting cRNA transcribed *in vitro* from the full-length cDNA clone into *Xenopus* oocytes, or by transfecting a mammalian cell line with a vector containing the cDNA linked to a suitable promoter.

2. Antibodies Against PN5.

The polypeptides of the invention are highly useful for the development of antibodies against PN5. Such antibodies can be used in affinity chromatography to purify recombinant sodium channel proteins or polypeptides, or they can be used as a research tool. For example, antibodies bound to a reporter molecule can be used in histochemical staining techniques to identify other tissues and cell types where PN5 are present, or they can be used to identify epitopic or functional regions of the sodium channel protein of the invention.

The antibodies can be monoclonal or polyclonal and can be prepared by techniques that are well known in the art. Polyclonal antibodies are prepared as follows: an immunogenic conjugate comprising PN5 or a fragment thereof, optionally linked to a carrier protein, is used to immunize a selected mammal such as a mouse, rabbit, goat, etc. Serum from the immunized mammal is collected and treated according to known procedures to separate the immunoglobulin fraction.

Monoclonal antibodies are prepared by standard hybridoma cell technology based on that reported by Kohler and Milstein in Nature 256, 495-497 (1975). Spleen cells are obtained from a host animal immunized with the PN5 protein or a fragment thereof, optionally linked to a carrier. Hybrid cells are formed by fusing these spleen cells with an appropriate myeloma cell line and cultured. The antibodies produced by the hybrid cells are screened for their ability to bind to expressed PN5 proteins.

A number of screening techniques well known in the art, such as, for example, forward or reverse enzyme-linked immunosorbent assay screening methods, may be employed. The hybrid cells producing such antibodies are then subjected to recloning and high dilution conditions in order to select a hybrid cell that secretes a homogeneous population of antibodies specific to either the PN5 protein.

In addition, antibodies can be raised by cloning and expressing nucleotide sequences or mutagenized versions thereof coding at least for the amino acid sequences required for specific binding of natural antibodies, and these expressed proteins used as the immunogen. Antibodies may include the complete immunoglobulin or a fragment thereof. Antibodies may be linked to a reporter group such as is described above with reference to polynucleotides.

Example 10 illustrates practice of producing an antibody.

3. Therapeutic Targets for Compounds to Treat Disorders and Assays Thereof.

The present invention also includes the use of the novel voltage-gated, preferably TTX-resistant, sodium channel α-subunit as a therapeutic target for compounds to treat disorders of the nervous system based on the RT-PCR localization data. The disorders include, but are not limited to, epilepsy, stroke injury, brain injury, diabetic neuropathy, traumatic injury, chronic neuropathic pain, and AIDS-associated neuropathy.

4. Designing Therapeutics based on Inhibiting PN5 and assays thereof.

This invention is also directed to inhibiting the activity of PN5 in brain, spinal cord, DRG, nodose ganglia, and superior cervical ganglia tissues. However, it is to be understood that further studies may reveal that PN5 is present in other tissues, and as such, those tissues can also be targeted areas. For example, the detection of PN5 mRNA in nodose ganglia suggests that PN5 may conduct TTX-resistant sodium currents in this and other sensory ganglia of the nervous system.

In addition, it has been found that proteins not normally expressed in certain tissues are expressed in a disease state. Therefore, this invention is intended to encompass the inhibition

of PN5 in tissues and cell types where the protein is normally expressed, and in those tissues and cell types where the protein is only expressed during a disease state.

For example, it is believed that TTX-resistant sodium channels play a key role in transmitting nerve impulses relating to sensory inputs such as pain and pressure. This information will facilitate the design of therapeutics that can be targeted to a specific area such as peripheral nerve tissue.

The recombinant protein of the present invention can be used to screen for potential therapeutics that have the ability to inhibit the sodium channel of interest. In particular, it would be useful to inhibit selectively the function of sodium channels in peripheral nerve tissues responsible for transmitting pain and pressure signals without simultaneously affecting the function of sodium channels in other tissues such as heart and muscle. Such selectivity would allow for the treatment of pain without causing side effects due to cardiac or neuromuscular complications. Therefore, it would be useful to have DNA sequences coding for sodium channels that are selectively expressed in peripheral nerve tissue.

5. Pain Reliever.

Sodium channels in peripheral nerve tissue play a large role in the transmission of nerve impulses, and therefore are instrumental in understanding neuropathic pain transmission. Neuropathic pain falls into two components: allodynia, where a normally non-painful stimulus becomes painful, and hyperalgesia, where a usually normal painful stimulus becomes extremely painful.

In tissue localization studies, PN5 mRNA maps small and medium neurons of DRG. PN5 mRNA is also present in brain and spinal cord. Inhibiting its activities may help prevent ailments such as headaches and migraines. The ability to inhibit the activity of these sodium channels, i.e., reduce the conduction of nerve impulses, will affect the nerve's ability to transmit pain impulses. Selective inhibition of sodium channels in sensory neurons such as DRG will allow the blockage of pain impulses without complicating side effects caused by inhibition of sodium channels in other tissues such as brain and heart. In addition, certain

diseases are caused by sodium channels that produce impulses at an extremely high frequency. The ability to reduce the activity of the channel can then eliminate or alleviate the disease. Accordingly, potential therapeutic compounds can be screened by methods well known in the art to discover whether they can inhibit the activity of the recombinant sodium channel of the invention. Barram, M. *et al.*, Naun-Schmiedeberg's Archives of Pharmacology 347, 125-132 (1993) and McNeal, E.T. *et al.*, J. Med. Chem. 28, 381-388 (1985). For similar studies with the acetyl choline receptor, see, Claudio *et al.*, Science 238, 1688-1694 (1987).

For example, pain can be alleviated by inhibiting the activity of the novel preferably TTX-resistant sodium channel comprising administering a therapeutically effective amount of a compound having an IC_{50} approximately 10 μM or less, preferably ≤1 μM. Potential therapeutic compounds are identified based on their ability to inhibit the activity of PN5. Therefore, the aforementioned assay can be used to identify compounds having a therapeutically effective IC_{50}.

The term „IC_{50}" refers to the concentration of a compound that is required to inhibit by 50% the activity of expressed PN5 when activity is measured by electrophysiology, flux assays, and toxin-binding assays, as mentioned above.

6. Diagnostic Assays.

The basic molecular biology techniques employed in accomplishing features of this invention, such as RNA, DNA and plasmid isolation, restriction enzyme digestion, preparation and probing of a cDNA library, sequencing clones, constructing expression vectors, transforming cells, maintaining and growing cell cultures, and other general techniques are well known in the art, and descriptions of such techniques can be found in general laboratory manuals such as Molecular Cloning: A Laboratory Manual by Sambrook *et al.* (Cold Spring Harbor Laboratory Press, 2nd edition, 1989).

For example, the polynucleotides of the invention can be bound to a „reporter molecule" to form a polynucleotide probe useful for Northern and Southern blot analysis and *in situ* hybridizations.

The term "reporter molecule" refers to a chemical entity capable of being detected by a suitable detection means, including, but not limited to, spectrophotometric, chemiluminescent, immunochemical, or radiochemical means. The polynucleotides of this invention can be conjugated to a reporter molecule by techniques well known in the art. Typically the reporter molecule contains a functional group suitable for attachment to or incorporation into the polynucleotide. The functional groups suitable for attaching the reporter group are usually activated esters or alkylating agents. Details of techniques for attaching reporter groups are well known in the art. See, for example, Matthews, J.A., Batki, A., Hynds, C., and Kricka, L.J., Anal. Biochem. 151, 205-209 (1985) and Engelhardt *et al.*, European Patent Application No. 0302175.

Accordingly, the following Examples are merely illustrative of the techniques by which the invention can be practiced.

Abbreviations

The following abbreviations are used throughout the Examples and have each of the respective meanings defined below.

BSA: bovine serum albumin

Denhardt's solution: 0.02% BSA, 0.02% polyvinyl-pyrrolidone, 0.02% Ficoll (0.1 g BSA, 0.1 g Ficoll and 0.1 g polyvinylpyrrolidone per 500 ml)

DRG: dorsal root ganglia

EDTA: Ethylenediaminetetraacetic acid, tetrasodium salt

MEN: 20 mM MOPS, 1 mM EDTA, 5 mM sodium acetate, pH 7.0

MOPS: 3-(N-morpholino)propanesulfonic acid (Sigma Chemical Company)

PN5: peripheral nerve sodium channel 5

PNS: peripheral nervous system

SDS: sodium dodecyl sulfate

SSC: 150 mM NaCl, 15 mM sodium citrate, pH 7.0

SSPE: 80 mM NaCl, 10 mM sodium phosphate, 1 mM ethylenediaminetetraacetate, pH 8.0

TEV: two electrode voltage clamp

TTX: tetrodotoxin (Sigma Chemical Company)

EXAMPLES

The following Examples illustrate practice of the invention.

Materials

The plasmid pBK-CMV was obtained from Stratagene (La Jolla, CA); the plasmid pBSTA is described by Goldin *et al.*, in Methods in Enzymology (Rudy & Iverson, eds.) 207, 279-297; the plasmid pCIneo was obtained from Promega (Madison, WI); and the plasmid pCRII was obtained from Invitrogen (Carlsbad, CA).

The oocyte expression vector plasmid pBSTAcIIr was constructed from pBSTA by insertion of a synthetic oligonucleotide linker; plasmid pKK232-8 was obtained from Pharmacia Biotech (Piscataway, NJ); plasmid pCRII was obtained from Invitrogen, San Diego, CA. Competent *E. coli* cell lines STBL2™ and SURE® were obtained from Gibco/BRL and Stratagene, respectively.

EXAMPLE 1

OBTAINING RNA FROM RAT DRG, BRAIN AND SPINAL CORD

Lumbar DRG No. 4 and No. 5 (L4 and L5) brain and spinal cord were removed from anesthetized adult male Sprague-Dawley rats under a dissecting microscope. The tissues were frozen in dry ice and homogenized with a Polytron homogenizer; the RNA was extracted by the guanidine isothiocyanate procedure (see Chomczynksi *et al.*, Anal. Biochemistry 162, 156-159 (1987)). Total RNA (5 μg of each sample) was dissolved in MEN buffer containing 50% formamide, 6.6% formaldehyde and denatured at 65°C for 5-10 min. The RNA was electrophoresed through a 0.8% agarose gel containing 8.3% formaldehyde in MEN buffer. The electrode buffer was MEN buffer containing 3.7% formaldehyde; the gel was run at 50 V for 12-18 hours.

Size markers, including ribosomal 18S and 28S RNAs and RNA markers (GIBCO BRL), were run in parallel lanes of the gel. Their positions were determined by staining the excised lane with ethidium bromide (0.5 μg/ml) followed by photography under UV light.

After electrophoresis, the gel was rinsed in 2xSSC and the RNA was transferred to a Duralose membrane (Stratagene) with 2OxSSC by capillary action; the membrane was baked under vacuum at 80°C for 1 hour.

EXAMPLE 2

PROBE FROM RAT BRAIN IIA

A ^{32}P-labeled cRNA probe complementary to nucleotides 4637-5868 of the rat brain IIA sodium channel α-subunit sequence was synthesized *in vitro* with T7 RNA polymerase (Pharmacia) using pEAF8 template DNA, (Noda *et al.*, Nature 320, 188-192 (1986)) that had been linearized with BstEII.

Protocols for each procedure mentioned above can be found in Molecular Cloning: A Laboratory Manual by Sambrook *et al.* (Cold Spring Harbor Laboratory Press, 2nd edition, 1989).

EXAMPLE 3

HYBRIDIZATION OF RNA WITH THE PROBE FROM RAT BRAIN IIA

The membrane of Example 1 was prehybridized in 50% formamide, 5xSSC, 50 mM sodium phosphate, pH 7.1, 1x Denhardt's solution, 0.5% SDS, and sheared, heat-denatured salmon sperm DNA (1 mg/ml) for 16 hours at 42°C. The membrane was hybridized in 50% formamide, 5xSSC, 50 mM sodium phosphate, pH 7.1, 1x Denhardt's solution, 0.5% SDS, and sheared, heat-denatured salmon sperm DNA (200 μg/ml) with the ^{32}P-labeled cRNA probe (ca. 1-3x10^6 cpm/ml) described in Example 2 for 18 hours at 42°C.

The membrane was rinsed with 2xSSC, 0.1% SDS at room temperature for 20 min. and then washed sequentially with: 2xSSC, 0.1%- SDS at 55°C for 30 min., 0.2xSSC, 0.1% SDS at 65°C for 30 min., 0.2xSSC, 0.1% SDS at 70°C for 30 min., and 0.2xSSC, 0.1% SDS, 0.1% sodium pyrophosphate at 70°C for 20 min. The filter was exposed against Kodak X-omat AR film at -80°C with intensifying screens for up to 2 weeks.

The pEAF8 probe hybridized to mRNAs in the DRG sample with sizes of 11 kb, 9.5 kb, 7.3 kb, and 6.5 kb, estimated on the basis of their positions relative to the standards.

EXAMPLE 4

NOVEL SODIUM CHANNEL DOMAIN IV PROBE

The probe was obtained as follows: RT-PCR was performed on RNA isolated from rat DRG using degenerate oligonucleotide primers that were designed based on the homologies between known sodium channels in domain IV. The domain IV products were cloned into a plasmid vector, transformed into E. coli and single colonies isolated. The domain IV specific PCR products obtained from several of these colonies were individually sequenced. Cloned novel domain IV sequence was as follows (SEQ ID NO: 4):

```
1    CTCAACATGG TTACGATGAT GGTGGAGACC GACGAGCAGG GCGAGGAGAA
51   GACGAAGGTT CTGGGCAGAA TCAACCAGTT CTTTGTGGCC GTCTTCACGG
101  GCGAGTGTGT GATGAAGATG TTCGCCCTGC GACAGTACTA TTTCACCAAC
151  GGCTGGAACG TGTTCGACTT CATAGTGGTG ATCCTGTCCA TTGGGAGTCT
201  GCTGTTTCT  GCAATCCTTA AGTCACTGGA AAACTACTTC TCCCCGACGC
251  TCTTCCGGGT CATCCGTCTG GCCAGGATCG GCCGCATCCT CAGGCTGATC
301  CGAGCAGCCA AGGGGATTCG CACGCTGCTC TTCGCCCTCA TGATGTCCCT
351  GCCCGCCCTC TTCAACATCG GCCTCCTCCT CTTCCTCGTC ATGTTCATCT
401  ACTCCATCTT CGGCATGGCC AGCTTCGCTA ACGTCGTGGA CGAGGCCGGC
451  ATCGACGACA TGTTCAACTT CAAGACCTTT GGCAACAGCA TGCTGTGCCT
501  GTTCCAGATC ACCACCTCGG CCGGCTGGGA CGGCCTCCTC AGCCCCATCC
551  TCAACACGGG GCCTCCCTAC TGCGACCCCA ACCTGCCCAA CAGCAACGGC
601  TCCCGGGGGA ACTGCGGGAG CCCGGCGGTG GGCATCATCT TCTTCACCAC
651  CTACATCATC ATCTCCTTCC TCATCGTGGT CAACATGTAT ATCGCAGTCA
701  TC
```

This sequence was labeled with ^{32}P by random priming.

EXAMPLE 5

HYBRIDIZATION OF RNA WITH THE NOVEL SODIUM CHANNEL 3'-UTR PROBE

A Northern blot was prepared with 10μg total RNA from rat brain, spinal cord, and DRG. The blot was hybridized with a cRNA probe from the 3'-UTR. The 3'-UTR was cloned into pSP 73 vector, the cRNA transcribed using a Trans Probe T kit (Pharmacia Biotech) and ^{32}P UTP. The blot was prehybridized for 2 hours at 65°C in a solution containing 5XSSC, 1X Denhardt's solution, 0.5% SDS, 50mM sodium phosphate, pH 7.1, salmon sperm DNA (1mg/ml) and 50% formamide. Hybridization was conducted at 45°C for 18 hours in the above solution except that the salmon sperm DNA was included at a concentration of 200μg/ml and the ^{32}P-labeled probe was added at $7.5x10^5$ cpm.ml solution. The blot was subsequently washed three times at 2XSSC and 0.1% SDS at room temperature, once with 0.2XSSC and 0.1% SDS at 65°C for 20 min., and once with 0.2XSSC, 0.1% SDS and 0.1% sodium pyrophosphate at 65°C for 20 min. The blot was analyzed on a PhosphoImager (BioRad) after an exposure of 2 days. The results indicated that there was a ~6.5kb band signal present in brain only in the lane containing RNA from DRG. Because of the lower abundance of PN5 mRNA, as evidenced by the RT-PCR experiment, the 6.5kb band was not detectable in brain and spinal cord.

EXAMPLE 6

CONSTRUCTION & SCREENING OF cDNA LIBRARY FROM RAT DRG

An EcoRI-adapted cDNA library was prepared from normal adult male Sprague-Dawley rat DRG poly(A)+ RNA using the SuperScript Choice System (GIBCO BRL). cDNA (>4 kb) was selected by sucrose gradient fractionation as described by Kieffer, Gene 109, 115-119 (1991). The cDNA was then ligated into the Zap Express vector (Stratagene), and packaged with the Gigapack II XL lambda packaging extract (Stratagene). Similarly, a >2kb DRG cDNA library was synthesized.

Phage ($3.5x10^5$) were screened by filter hybridization with a ^{32}P-labeled probe (rBIIa, bases 4637-5868 as follows of Auld *et al.*, Neuron 1, 449-461 (1988)). Filters were hybridized in 50% formamide, 5X SSPE, 5X Denhardt's solution, 0.5% SDS, 250µg/ml sheared, denatured salmon sperm DNA, and 50 mM sodium phosphate at 42°C and washed in 0.5X SSC/0.1% SDS at 50°C.

Southern blots of EcoRI-digested plasmids were hybridized with the ^{32}P-labeled DNA probe, (SEQ ID NO: 4). The filters were then hybridized in 50% formamide, 6X SSC, 5X Denhardt's solution, 0.5%, SDS, and 100µg/ml sheared, denatured salmon sperm DNA at 42°C and were washed in 0.1X SSC/0.1% SDS at 65°C.

Positive clones were excised in vivo into pBK-CMV using the ExAssist/XLOLR system (Stratagene).

EXAMPLE 7

CLONES AND NUCLEOTIDE ANALYSIS

cDNA clones, 26.2 and 25.1 were isolated from the >4kb DRG cDNA library and clone 1.18 was isolated from the >2kb DRG cDNA library. By sequence analysis, 26.2 appeared to be a full-length cDNA encoding a novel sodium channel and 25.1 extended from domain II to the 3'-UTR. However, each had a deletion which truncated the coding region. Clone 1.18 had the 3'- untranslated region, in addition to the C-terminus of the deduced amino acid sequence of PN5. The construct in the expression vector, pBSTACIIr, consisted of sequences from 26.2 and 1.18.

PN5 homology to other known sodium channels was obtained using the GAP/Best Fit (GCG) program:

Channel	% Similarity	% Identity
PN3a	71	54
hPN3	71	55
PN4	71	53
PN4a	71	53

PN1	72	55
rat brain type I	72	55
rat brain type II	71	54
rat brain type III	71	54
rat cardiac channel	73	56
rat skeletal muscle channel	71	53

Stabilizing the PN5 full length cDNA

A. Media, *E. coli* cell lines, and growth conditions:

Growth of fragments of PN5 could be accomplished under standard conditions; however growth of plasmids containing full length constructs of PN5 (in pCIneo, pBSTAcIIr, and other vectors) could not be accomplished without use of special growth media, conditions, and *E. coli* strains. The following proved to be optimal: (1) use of *E. coli* STBL2™ for primary transformation following ligation reactions and for large scale culturing; (2) solid media was 1/2x FM (see below) plus 1x LB (Tryptone, 1%, Yeast Extract, 0.5%, NaCl, 0.5%), plus 15g/L agar, or 1xFM plus 1/2x LB; (3) liquid media optimally was 1x FM plus 1/2x LB; (4) carbenicillin, 100μg/ml, was used for all media, as it is metabolized less rapidly than ampicillin; (5) temperature for growth should be no greater than 30°C, usually 24-26°C; this necessitated longer growth periods than normally employed, from 24 to 72 hours.

2x Freezing Medium (2xFM):

K2HP04	12.6g
Na3Citrate	0.9g
MgSO4.7H20	0.18g
(NH4)2SO4	1.8g
KH2PO4	3.6g
Glycerol	88g
H20	qs to IL

2x FM and the remaining media components are prepared separately, sterilized by autoclaving, cooled to at least 60°C, and added together to form the final medium. Carbenicillin is prepared

at 25mg/ml H20 and sterilized by filtration. 2x FM was first described for preparation of frozen stocks of bacterial cells (Practical Methods in Molecular Biology, Schleif, R.F. and Wensink, P.C., Springer-Verlag, New York (1981) pp. 201-202).

B. Expression Vectors

In order to provide for increased stability of the full length cDNA, the oocyte expression vector pBSTAcIIr was modified to reduce plasmid copy number when grown in *E. coli* and to reduce possible read-through transcription from vector sequences that might result in toxic cryptic expression of PN5 protein, Brosius J., Gene 27, 151-160(1984). pBSTAcIIr was digested with PvuII. The 755 bp fragment containing the T7 promoter, ß-globin 5'UTR, the multiple cloning site, ß-globin 3'UTR, and T3 promoter was ligated to the 3.6 kb fragment containing the replication origin, ampicillin resistance gene, rrnBT_1 and rrnBT_1T_2 transcription terminators from pKK232-8, which had been fully digested with SmaI and partially digested with PvuII and treated with shrimp intestinal phosphatase to prevent self ligation. The resulting plasmid in which the orientation of the pBSTA fragment is such that the T7 promoter is proximal to the rrnBT_1 terminator was identified by restriction mapping and named pHQ8. As is the case with pBSTA, the direction of transcription of the ampicillin resistance gene and replication origin of pHQ8 is opposite to that of the gene expression cassette, and the presence of the rrnB T_1 terminator should reduce any remaining read-through from the vector into the T7 promoter driven expression cassette.

C. Assembly of full length cDNA for expression

Since pBK-CMV.26.2 had a 58 bp deletion (corresponding to bp 4346 to 4403 of SEQ ID NO: 1) and the sequnce of pBK-CMV.1.18 begins at bp 4180 of SEQ ID NO: 1, pBK-CMV.1.18 could be used to „repair" pBK-CMV.26.2. A strategy was developed to assemble a full length cDNA from clones pBK-CMV.26.2 and pBK-CMV.1.18 in three sections, truncating the 5' and 3' UTRs and introducing unique restriction sites at the 5' and 3' ends in the process. The 5' end

was generated by PCR from 26.2, truncating the 5' UTR by incorporating a SalI site just upstream of the start codon. The central section was a restriction fragment from 26.2. The 3' end was prepared by overlap PCR from both 26.2 and 1.18 and incorporating an XbaI site just down stream of the stop codon. These sections were digested at unique restriction sites and assembled in pBSTAcIIr. Although this construct appeared to have a correct sequence, upon recloning as a SalI to XbaI fragment into pCIneo, two type of isolates were found, one with a deletion and one with an 8 bp insertion. Reexamination of the pBSTAcIIr clone showed the sequence was „mixed" in this region, so that the clone must have rearranged. The 8 bp insertion was found as a repeat of one of the members of an 8 bp duplication in the native sequence, forming a triple 8 bp repeat in the rearranged isolate. Numerous cloning attempts inevitably gave rise to this rearrangement. Overlap PCR was used to introduce silent mutations into one of the 8 bp repeats, and a fragment containing this region was included when the PN5 coding region was assembled into HQ8, the low-copy number version of pBSTAcIIr, to give plasmid HR-1. This sequence proved to be stable (see Figures 5A-E, SEQ ID NO: 5).

The 5' end fragment was prepared by PCR using pBK-CMV.26.2 DNA as template and primers 4999 (CTTGGTCGACTCTAGA*TCAGGGTGAAGATGGAGGAG*; SalI site underlined, PN5 homology in italics, corresponding to bp 58-77 of SEQ ID NO: 1, initiation codon in bold) and 4927 (GGGTTCAATGTGGTTTTATCT, corresponding to bp 1067 to 1047 of SEQ ID NO: 1), followed by gel purification, digestion with SalI and KpnI (KpnI site at pb 1003-1008, SEQ ID NO: 1), and gel purification.

The central 3.1 kb fragment was prepared by digestion of pBK-CMV.26.2 DNA with KpnI and AatII (AatII site at 4133-4138), followed by gel purification.

The 3' end fragment was prepared as follows: PCR using primers 4837 (TCTGGGAAGTTTGGAAG, corresponding to bp 3613 to 3629 of SEQ ID NO: 1) and 4931

(GACCACGAAGGCTATGTTGAGG, corresponding to bp 4239 to 4218 of SEQ ID NO: 1) on pBK-CMV.26.2 DNA as template gave a fragment of 0.6 kb. PCR using primers 4930 (CCTCAACATAGCCTTCGTGGTC, corresponding to bp 4218 to 4239 of SEQ ID NO: 1) and 4929 (GTCTTCTAGA*ATGAGGGTTCAGTCATTGTG*, XbaI site underlined, PN5 homology in italics, corresponding to pb 5386 to 5365 of SEQ ID NO: 1, stop codon in bold) on pBK-CMV.1.18 DNA as template gave a fragment of 1.2 kb, introducing a XbaI site 7 bp from the stop codon. Thus the 3' end of the 4837-4931 fragment exactly complements the 5' end of the 4930-4929 fragment. These two fragments were gel purified and a fraction of each combined as template in a PCR reaction using primers 4928 (CAAGCCTTTGTGTTCGAC, corresponding to bp 4084 to 4101 of SEQ ID NO: 1) and 4929, to give a fragment of 1.3 kb. This fragment was gel purifed, digested with AatII and XbaI, and the 1.2 kb fragment gel purified.

The 3' end fragment was cloned into AatII and XbaI digested pBSTAcIIr. One isloate was digested with SalI and KpnI and ligated to the 5' end fragment. The resulting plasmid, after sequence verification, was digested with KpnI and AatII and ligated to the central 3.1 kb fragment, to form pBSTAcIIr.PN5(clone 21). pBSTAcIIr.PN5 (clone 21) was digested with SalI and XbaI to release the 5.3 kb PN5 fragment which was cloned into SalI and XbaI digested pCIneoII. Multiple isolates were found, of which GPII-1, which was completely sequenced, was typical and contained an 8 bp insert. This CAGAAGAA, after pb 3994 of SEQ ID NO: 1, converted the direct repeat of this sequence at this location into a triple direct repeat, causing a shift in the reading frame. In an attempt to repair this defect, pBSTAcIIr. PN5 (clone 21) was digested with NheI (bp 2538-2543 SEQ ID NO: 1) and XhoI (bp 4828-4833, SEQ ID NO: 1) to give a 6.2 kb fragment and with AatII and XhoI to give a 0.7 kb fragment which were ligated to the 1.6 kp fragment resulting from digestion of pBK-CMV.26.2 with AatII and NheI. Although no isolates were found which were completely correct, one isolate, HA-4, had only a single base

change, deletion of the C at bp 4827 (SEQ ID NO: 1) adjacent to the XhoI site.

In order to prevent the 8 bp insertion rearrangement from occurring, three silent mutations were introduced in the 5' repeat, and two additional mutations in a string ot Ts would also be introduced, as shown below (bp 3982 to 4014, SEQ ID NO: 1; mutation sites underlined, 8 bp repeats in native sequence in italics):

native	GAC	ATT	TTT	ATG	ACA	*GAA*	*GAA*	*CAG*	*AAG*	*AAA*	TAT
	Asp	Ile	Phe	Met	Thr	Glu	Glu	Gln	Lys	Lys	Tyr
mutant	GAC	ATC	TTC	ATG	ACT	GAG	GAG	CAG	AAG	AAA	TAT

As isolate HA-4 had the native direct repeat sequence (as opposed to e.g. pBSTAcIIr.PN5 (clone 21)) and the region near the XhoI site defect would not be involved, it was used as template DNA for the following PCR reactions. Primer P5-3716S (CCGAAGCCAATGTAACATTAGTAATTACTCGTG, corresponding to pb 3684 to 3716, SEQ ID NO: 1) was paired with primer P5-3969AS (GCTCCTCAGTCATGAAGATGTCTTGGCCACCTAAC, correspoind to bp 4003 to 3969, SEQ ID NO: 1, mutated bases are underlined) to give a 320 bp product. Primer P5-4017S (GGCCAAGACATCTTCATGACTGAGGAGCAGAAGAAATATTAC, corresponding to bp 3976 to 4017, SEQ ID NO: 1; mutated bases are underlined) was paired with primer P5-4247AS (CTCAAAGCAAAGACTTTGATGAGACACTCTATGG, corresoinding to bp 4280 to 4247, SEQ ID NO: 1) to give a 305 bp product. The 3' end of the 320 bp fragment thus has a 28 bp exact match to the 5' end of the 305 bp fragment. The two bands were gel purified and a fraction of each combined in a new PCR reaction with primers P5-3716S and P5-4247AS to give a 597 bp product, which was T/A cloned into vector pCRII. Isolate HO-7 was found to have the desired sequence. A four-way ligation was performed to assemble the full-length, modified PN5:

the oocyte expression vector HQ-8 ws digested with SalI and XbaI to give a 4.4 kb vector fragment; GPII-1 was digested wtih SalI and MluI to give a 3.8 kb fragment containing the 5' half of PN5; HO-7 was digested with MluI (bp 3866 to 3871, SEQ ID NO: 1) and AatII to give a 0.3 kb fragment containing the mutant 8 bp repeat region of PN5; GPII-1 was digested with AatII and XbaI to give the remaining 1.3 kb 3' portion of PN5. A portion of the ligation reaction was transformed into *E. coli* Stable 2 cells. Of the 9.6 kb isolates containing all four fragments, HR-1 was sequenced and found to have the desired 5.4 kb sequence. These isolates grew well and showed no tendency to rearrange. The sequence of this engineered version of PN5 is shown in Figures 5A-E (SEQ ID NO: 5).

EXAMPLE 8

HUMAN PN5

An 856 bp clone (Figure 3A, SEQ ID No.: 3) has been isolated from a human dorsal root ganglia (DRG) cDNA library that is most closely related to rat PN5 with 79% identity for the amino acid sequence. The human PN5 sequence spans the region between IIIS1 and interdomain III/IV which includes the fast inactivation gate (i.e., IFM) that is located within interdomain III/IV.

The human DRG cDNA library was constructed from lumbar 4 and 5 DRG total RNA that was randomly primed. First strand cDNA was synthesized with SuperScript II reverse transcriptase (GIBCO BRL) and the second strand synthesis with T4 DNA polymerase. EcoRI adaptors were ligated to the ends of the double stranded cDNAs and the fragments cloned into the ZAP II vector (Stratagene). The library was screened with digoxigenin-labeled rat PN3, rat PN1 and human heart hH1 probes. Positive clones were sequenced and compared to known human and rat sodium channel sequences. Only the aforementioned clone was identified as human PN5 sequence.

Channel	% Similarity	% Identity
Human Brain (HBA)	76	69
Human Heart (hH1)	81	74

Human Atypical Heart	60	52
Human Skeletal Muscle	80	71
Human Neuroendocrine	78	71
Human PN3	77	70
Rat PN1	79	72
Rat PN3	78	71
Rat PN4	78	70
Rat PN5	86	79

Figure 3B compares the amino acid sequence of the hPN5 fragment with the rat PN5 amino acid sequence in the appropriate region.

EXAMPLE 9

TISSUE DISTRIBUTION BY RT-PCR

Brain, spinal cord, DRG, nodose ganglia, superior cervical ganglia, sciatic nerve, heart and skeletal muscle tissue were isolated from anesthetized, normal adult male Sprague-Dawley rats and were stored at -80°C. RNA was isolated from each tissue using RNAzol (Tel-Test, Inc.). Random-primed cDNA was reverse transcribed from 500ng of RNA from each tissue. The forward primer (CAGATTGTGTTCTCAGTACATTCC) and the reverse primer (CCAGGTGTCTAACGAATAAATAGG) were designed from the 3'-untranslated region to yield a 252 base pair fragment. The cycle parameters were: 94°C/2 min. (denaturation), 94°C/30 sec., 65°C/30 sec. and 72°C/1min. (35 cycles) and 72°C/4 min. The reaction products were analyzed on a 4% agarose gel.

A positive control and a no-template control were also included. cDNA from each tissue was also PCR amplified using primers specific for glyceraldehyde-3-phosphate dehydrogenase to demonstrate template viability, as described by Tso *et al.*, Nucleic Acid Res. 13, 2485-2502 (1985).

Tissue distribution profile of rPN5 by analysis of RNA from selected rat tissues by RT-PCR was as follows:

Tissue	RT-PCR (35 cycles)
Brain	+

Spinal cord	+
DRG	+++
Nodose ganglia	+++
Superior cervical ganglia	+
Sciatic nerve	-
Heart	-
Skeletal muscle	-
F11-untreated	+
F11-treated	+

PN5 was also detected after only 25 cycles (24 + 1) in the same five tissues as above in the same relative abundance.

EXAMPLE 10

ANTIBODIES

A synthetic peptide (26 amino acids in interdomain II and III - residues 977 to 1002) was conjugated to KLH and antibody raised in rabbits. The antiserum was subsequently affinity purified.

PN5 constitutes a subfamily of novel sodium channel genes; these genes are different from those detectable with other probes (e.g., PEAF8 and PN3 probes).

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: F. HOFFMANN-LA ROCHE AG
(B) STREET: Grenzacherstrasse 124
(C) CITY: Basle
(D) STATE: BS
(E) COUNTRY: Switzerland
(F) POSTAL CODE (ZIP): CH-4010
(G) TELEPHONE: 061-6884256
(H) TELEFAX: 061-6881395
(I) TELEX: 962292/965542 hlr ch

(ii) TITLE OF INVENTION: Nucleic Acid Encoding a Nervous Tissue Sodium Channel
(iii) NUMBER OF SEQUENCES: 5
(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release (1.0, Version (1.30
(v) CURRENT APPLICATION DATA
(A) APPLICATION NUMBER:
(B) FILING DATE:

(2) INFORMATION FOR SEQ ID NO:1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5908 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(vi) ORIGINAL SOURCE:
(A) ORGANISM: rat
(F) TISSUE TYPE: Dorsal root ganglia
(G) CELL TYPE: Peripheral nerve
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAAGTCACAG GAGTGTCTGT CAGCGAGAGG AAGAAGGGAG AGTTTACTGA GTGTCTTCTG 60

CCCCTCCTCA GGGTGAAGAT GGAGGAGAGG TACTACCCGG TGATCTTCCC GGACGAGCGG 120

AATTTCCGCC CCTTCACTTC CGACTCTCTG GCTGCCATAG AGAAGCGGAT TGCTATCCAA 180

AAGGAGAGGA AGAAGTCCAA AGACAAGGCG GCAGCTGAGC CCCAGCCTCG GCCTCAGCTT 240

GACCTAAAGG CCTCCAGGAA GTTACCTAAG CTTTATGGTG ACATTCCCCC TGAGCTTGTA 300

GCGAAGCCTC TGGAAGACCT GGACCCATTC TACAAAGACC ATAAGACATT CATGGTGTTG 360

AACAAGAAGA GAACAATTTA TCGCTTCAGC GCCAAGCGGG CCTTGTTCAT TCTGGGGCCT 420

TTTAATCCCC TCAGAAGCTT AATGATTCGT ATCTCTGTCC ATTCAGTCTT TAGCATGTTC 480

ATCATCTGCA CGGTGATCAT CAACTGTATG TTCATGGCGA ATTCTATGGA GAGAAGTTTC 540

GACAACGACA TTCCCGAATA CGTCTTCATT GGGATTTATA TTTTAGAAGC TGTGATTAAA 600

ATATTGGCAA GAGGCTTCAT TGTGGATGAG TTTTCCTTCC TCCGAGATCC GTGGAACTGG 660

CTGGACTTCA TTGTCATTGG AACAGCGATC GCAACTTGTT TTCCGGGCAG CCAAGTCAAT 720

CTTTCAGCTC TTCGTACCTT CCGAGTGTTC AGAGCTCTGA AGGCGATTTC AGTTATCTCA 780

GGTCTGAAGG TCATCGTAGG TGCCCTGCTG CGCTCGGTGA AGAAGCTGGT AGACGTGATG 840

GTCCTCACTC TCTTCTGCCT CAGCATCTTT GCCCTGGTCG GTCAGCAGCT GTTCATGGGA 900

ATTCTGAACC AGAAGTGTAT TAAGCACAAC TGTGGCCCCA ACCCTGCATC CAACAAGGAT 960

TGCTTTGAAA AGGAAAAAGA TAGCGAAGAC TTCATAATGT GTGGTACCTG GCTCGGCAGC 1020

AGACCCTGTC CCAATGGTTC TACGTGCGAT AAAACCACAT TGAACCCAGA CAATAATTAT 1080

ACAAAGTTTG ACAACTTTGG CTGGTCCTTT CTCGCCATGT TCCGGGTTAT GACTCAAGAC 1140

TCCTGGGAGA GGCTTTACCG ACAGATCCTG CGGACCTCTG GGATCTACTT TGTCTTCTTC 1200

TTCGTGGTGG TCATCTTCCT GGGCTCCTTC TACCTGCTTA ACCTAACCCT GGCTGTTGTC 1260

ACCATGGCTT ATGAAGAACA GAACAGAAAT GTAGCTGCTG AGACAGAGGC CAAGGAGAAA 1320

ATGTTTCAGG AAGCCCAGCA GCTGTTAAGG GAGGAGAAGG AGGCTCTGGT TGCCATGGGA 1380

ATTGACAGAA GTTCCCTTAA TTCCCTTCAA GCTTCATCCT TTTCCCCGAA GAAGAGGAAG 1440

TTTTTCGGTA GTAAGACAAG AAAGTCCTTC TTTATGAGAG GGTCCAAGAC GGCCCAAGCC 1500

TCAGCGTCTG ATTCAGAGGA CGATGCCTCT AAAAATCCAC AGCTCCTTGA GCAGACCAAA 1560

CGACTGTCCC AGAACTTGCC AGTGGATCTC TTTGATGAGC ACGTGGACCC CCTCCACAGG 1620

CAGAGAGCGC TGAGCGCTGT CAGTATCTTA ACCATCACCA TGCAGGAACA AGAAAAATTC 1680

CAGGAGCCTT GTTTCCCATG TGGGAAAAAT TTGGCCTCTA AGTACCTGGT GTGGGACTGT 1740

AGCCCTCAGT GGCTGTGCAT AAAGAAGGTC CTGCGGACCA TCATGACGGA TCCCTTTACT 1800

GAGCTGGCCA TCACCATCTG CATCATCATC AATACCGTTT TCTTAGCCGT GGAGCACCAC 1860

AACATGGATG ACAACTTAAA GACCATACTG AAAATAGGAA ACTGGGTTTT CACGGGAATT 1920

TTCATAGCGG AAATGTGTCT CAAGATCATC GCGCTCGACC CTTACCACTA CTTCCGGCAC 1980

GGCTGGAATG TTTTTGACAG CATCGTGGCC CTCCTGAGTC TCGCTGATGT GCTCTACAAC 2040

ACACTGTCTG ATAACAATAG GTCTTTCTTG GCTTCCCTCA GAGTGCTGAG GGTCTTCAAG 2100

TTAGCCAAAT CCTGGCCCAC GTTAAACACT CTCATTAAGA TCATCGGCCA CTCCGTGGGC 2160

GCGCTTGGAA ACCTGACTGT GGTCCTGACT ATCGTGGTCT TCATCTTTTC TGTGGTGGGC 2220

ATGCGGCTCT TCGGCACCAA GTTTAACAAG ACCGCCTACG CCACCCAGGA GCGGCCCAGG 2280

CGGCGCTGGC ACATGGATAA TTTCTACCAC TCCTTCCTGG TGGTGTTCCG CATCCTCTGT 2340

GGGGAATGGA TCGAGAACAT GTGGGGCTGC ATGCAGGATA TGGACGGCTC CCCGTTGTGC 2400

ATCATTGTCT TTGTCCTGAT AATGGTGATC GGGAAGCTTG TGGTGCTTAA CCTCTTCATT 2460

GCCTTGCTGC TCAATTCCTT CAGCAATGAG GAGAAGGATG GGAGCCTGGA AGGAGAGACC 2520

AGGAAAACCA AAGTGCAGCT AGCCCTGGAT CGGTTCCGCC GGGCCTTCTC CTTCATGCTG 2580

CACGCTCTTC AGAGTTTTTG TTGCAAGAAA TGCAGGAGGA AAAACTCGCC AAAGCCAAAA 2640

GAGACAACAG AAAGCTTTGC TGGTGAGAAT AAAGACTCAA TCCTCCCGGA TGCGAGGCCC 2700

TGGAAGGAGT ATGATACAGA CATGGCTTTG TACACTGGAC AGGCCGGGGC TCCGCTGGCC 2760

CCACTCGCAG AGGTAGAGGA CGATGTGGAA TATTGTGGTG AAGGCGGTGC CCTACCCACC 2820

TCACAACATA GTGCTGGAGT TCAGGCCGGT GACCTCCCTC CAGAGACCAA GCAGCTCACT 2880

AGCCCGGATG ACCAAGGGGT TGAAATGGAA GTATTTTCTG AAGAAGATCT GCATTTAAGC 2940

ATACAGAGTC CTCGAAAGAA GTCTGACGCA GTGAGCATGC TCTCGGAATG CAGCACAATT 3000

GACCTGAATG ATATCTTTAG AAATTTACAG AAAACAGTTT CCCCCAAAAA GCAGCCAGAT 3060

AGATGCTTTC CCAAGGGCCT TAGTTGTCAC TTTCTATGCC ACAAAACAGA CAAGAGAAAG 3120

TCCCCCTGGG TCCTGTGGTG GAACATTCGG AAAACCTGCT ACCAAATCGT GAAGCACAGC 3180

TGGTTTGAGA GTTTCATAAT CTTTGTTATT CTGCTGAGCA GTGGAGCGCT GATATTTGAA 3240

GATGTCAATC TCCCCAGCCG GCCCCAAGTT GAGAAATTAC TAAGGTGTAC CGATAATATT 3300

TTCACATTTA TTTTCCTCCT GGAAATGATC CTGAAGTGGG TGGCCTTTGG ATTCCGGAGG 3360

TATTTCACCA GTGCCTGGTG CTGGCTTGAT TTCCTCATTG TGGTGGTGTC TGTGCTCAGT 3420

CTCATGAATC TACCAAGCTT GAAGTCCTTC CGGACTCTGC GGGCCCTGAG ACCTCTGCGG 3480

GCGCTGTCCC AGTTTGAAGG AATGAAGGTT GTCGTCTACG CCCTGATCAG CGCCATACCT 3540

GCCATTCTCA ATGTCTTGCT GGTCTGCCTC ATTTTCTGGC TCGTATTTTG TATCTTGGGA 3600

GTAAATTTAT TTTCTGGGAA GTTTGGAAGG TGCATTAACG GGACAGACAT AAATATGTAT 3660

TTGGATTTTA CCGAAGTTCC GAACCGAAGC CAATGTAACA TTAGTAATTA CTCGTGGAAG 3720

GTCCCGCAGG TCAACTTTGA CAACGTGGGG AATGCCTATC TCGCCCTGCT GCAAGTGGCA 3780

ACCTATAAGG GCTGGCTGGA AATCATGAAT GCTGCTGTCG ATTCCAGAGA GAAAGACGAG 3840

CAGCCGGACT TTGAGGCGAA CCTCTACGCG TATCTCTACT TTGTGGTTTT TATCATCTTC 3900

GGCTCCTTCT TTACCCTGAA CCTCTTTATC GGTGTTATTA TTGACAACTT CAATCAGCAG 3960

CAGAAAAAGT TAGGTGGCCA AGACATTTTT ATGACAGAAG AACAGAAGAA ATATTACAAT 4020

GCAATGAAAA AGTTAGGAAC CAAGAAACCT CAAAAGCCCA TCCCAAGGCC CCTGAACAAA 4080

TGTCAAGCCT TTGTGTTCGA CCTGGTCACA AGCCAGGTCT TTGACGTCAT CATTCTGGGT 4140

CTTATTGTCT TAAATATGAT TATCATGATG GCTGAATCTG CCGACCAGCC CAAAGATGTG 4200

AAGAAAACCT TTGATATCCT CAACATAGCC TTCGTGGTCA TCTTTACCAT AGAGTGTCTC 4260

ATCAAAGTCT TTGCTTTGAG GCAACACTAC TTCACCAATG GCTGGAACTT ATTTGATTGT 4320

GTGGTCGTGG TTCTTTCTAT CATTAGTACC CTGGTTTCCC GCTTGGAGGA CAGTGACATT 4380

TCTTTCCCGC CCACGCTCTT CAGAGTCGTC CGCTTGGCTC GGATTGGTCG AATCCTCAGG 4440

CTGGTCCGGG CTGCCCGGGG AATCAGGACC CTCCTCTTTG CTTTGATGAT GTCTCTCCCC 4500

TCTCTCTTCA ACATCGGTCT GCTGCTCTTC CTGGTGATGT TCATTTACGC CATCTTTGGG 4560

ATGAGCTGGT TTTCCAAAGT GAAGAAGGGC TCCGGGATCG ACGACATCTT CAACTTCGAG 4620

ACCTTTACGG GCAGCATGCT GTGCCTCTTC CAGATAACCA CTTCGGCTGG CTGGGATACC 4680

CTCCTCAACC CCATGCTGGA GGCAAAAGAA CACTGCAACT CCTCCTCCCA AGACAGCTGT 4740

CAGCAGCCGC AGATAGCCGT CGTCTACTTC GTCAGTTACA TCATCATCTC CTTCCTCATC 4800

GTGGTCAACA TGTACATCGC TGTGATCCTC GAGAACTTCA ACACAGCCAC GGAGGAGAGC 4860

GAGGACCCTC TGGGAGAGGA CGACTTTGAA ATCTTCTATG AGGTCTGGGA GAAGTTTGAC 4920

CCCGAGGCGT CGCAGTTCAT CCAGTATTCG GCCCTCTCTG ACTTTGCGGA CGCCCTGCCG 4980

GAGCCGTTGC GTGTGGCCAA GCCGAATAAG TTTCAGTTTC TAGTGATGGA CTTGCCCATG 5040

GTGATGGGCG ACCGCCTCCA TTGCATGGAT GTTCTCTTTG CTTTCACTAC CAGGGTCCTC 5100

GGGGACTCCA GCGGCTTGGA TACCATGAAA ACCATGATGG AGGAGAAGTT TATGGAGGCC 5160

AACCCTTTTA AGAAGCTCTA CGAGCCCATA GTCACCACCA CCAAGAGGAA GGAGGAGGAG 5220

CAAGGCGCCG CCGTCATCCA GAGGGCCTAC CGGAAACACA TGGAGAAGAT GGTCAAACTG 5280

AGGCTGAAGG ACAGGTCAAG TTCATCGCAC CAGGTGTTTT GCAATGGAGA CTTGTCCAGC 5340

TTGGATGTGG CCAAGGTCAA GGTTCACAAT GACTGAACCC TCATCTCCAC CCCTACCTCA 5400

CTGCCTCACA GCTTAGCCTC CAGCCTCTGG CGAGCAGGCG GCAGACTCAC TGAACACAGG 5460

CCGTTCGATC TGTGTTTTTG GCTGAACGAG GTGACAGGTT GGCGTCCATT TTTAAATGAC 5520

TCTTGGAAAG ATTTCATGTA GAGAGATGTT AGAAGGGACT GCAAAGGACA CCGACCATAA 5580

CGGAAGGCCT GGAGGACAGT CCAACTTACA TAAAGATGAG AAACAAGAAG GAAAGATCCC 5640

AGGAAAACTT CAGATTGTGT TCTCAGTACA TCCCCCAATG TGTCTGTTCG GTGTTTTGAG 5700

TATGTGACCT GCCACATGTA GCTCTTTTTT GCATGTACGT CAAAACCCTG CAGTAAGTTG 5760

ATAGCTTGCT ACGGGTGTTC CTACCAGCAT CACAGAATTG GGTGTATGAC TCAAACCTAA 5820

AAGCATGACT CTGACTTGTC AGTCAGCACC CCGACTTTCA GACGCTCCAA TCTCTGTCCC 5880

AGGTGTCTAA CGAATAAATA GGTAAAAG 5908

(3) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1765 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: rat

(F) TISSUE TYPE: dorsal root ganglia

(G) CELL TYPE: peripheral nerve

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Glu Glu Arg Tyr Tyr Pro Val Ile Phe Pro Asp Glu Arg Asn Phe
1               5                  10                  15
Arg Pro Phe Thr Ser Asp Ser Leu Ala Ala Ile Glu Lys Arg Ile Ala
            20                  25                  30
Ile Gln Lys Glu Arg Lys Lys Ser Lys Asp Lys Ala Ala Ala Glu Pro
        35                  40                  45
Gln Pro Arg Pro Gln Leu Asp Leu Lys Ala Ser Arg Lys Leu Pro Lys
    50                  55                  60
Leu Tyr Gly Asp Ile Pro Pro Glu Leu Val Ala Lys Pro Leu Glu Asp
65                  70                  75                  80
Leu Asp Pro Phe Tyr Lys Asp His Lys Thr Phe Met Val Leu Asn Lys
                85                  90                  95
Lys Arg Thr Ile Tyr Arg Phe Ser Ala Lys Arg Ala Leu Phe Ile Leu
            100                 105                 110
Gly Pro Phe Asn Pro Leu Arg Ser Leu Met Ile Arg Ile Ser Val His
        115                 120                 125
Ser Val Phe Ser Met Phe Ile Ile Cys Thr Val Ile Ile Asn Cys Met
    130                 135                 140
Phe Met Ala Asn Ser Met Glu Arg Ser Phe Asp Asn Asp Ile Pro Glu
145                 150                 155                 160
Tyr Val Phe Ile Gly Ile Tyr Ile Leu Glu Ala Val Ile Lys Ile Leu
                165                 170                 175
```

```
Ala Arg Gly Phe Ile Val Asp Glu Phe Ser Phe Leu Arg Asp Pro Trp
            180                 185                 190
Asn Trp Leu Asp Phe Ile Val Ile Gly Thr Ala Ile Ala Thr Cys Phe
        195                 200                 205
Pro Gly Ser Gln Val Asn Leu Ser Ala Leu Arg Thr Phe Arg Val Phe
    210                 215                 220
Arg Ala Leu Lys Ala Ile Ser Val Ile Ser Gly Leu Lys Val Ile Val
225                 230                 235                 240
Gly Ala Leu Leu Arg Ser Val Lys Lys Leu Val Asp Val Met Val Leu
                245                 250                 255
Thr Leu Phe Cys Leu Ser Ile Phe Ala Leu Val Gly Gln Gln Leu Phe
            260                 265                 270
Met Gly Ile Leu Asn Gln Lys Cys Ile Lys His Asn Cys Gly Pro Asn
        275                 280                 285
Pro Ala Ser Asn Lys Asp Cys Phe Glu Lys Glu Lys Asp Ser Glu Asp
    290                 295                 300
Phe Ile Met Cys Gly Thr Trp Leu Gly Ser Arg Pro Cys Pro Asn Gly
305                 310                 315                 320
Ser Thr Cys Asp Lys Thr Thr Leu Asn Pro Asp Asn Asn Tyr Thr Lys
                325                 330                 335
Phe Asp Asn Phe Gly Trp Ser Phe Leu Ala Met Phe Arg Val Met Thr
            340                 345                 350
Gln Asp Ser Trp Glu Arg Leu Tyr Arg Gln Ile Leu Arg Thr Ser Gly
        355                 360                 365
Ile Tyr Phe Val Phe Phe Phe Val Val Val Ile Phe Leu Gly Ser Phe
    370                 375                 380
Tyr Leu Leu Asn Leu Thr Leu Ala Val Val Thr Met Ala Tyr Glu Glu
385                 390                 395                 400
Gln Asn Arg Asn Val Ala Ala Glu Thr Glu Ala Lys Glu Lys Met Phe
                405                 410                 415
Gln Glu Ala Gln Gln Leu Leu Arg Glu Glu Lys Glu Ala Leu Val Ala
            420                 425                 430
Met Gly Ile Asp Arg Ser Ser Leu Asn Ser Leu Gln Ala Ser Ser Phe
        435                 440                 445
Ser Pro Lys Lys Arg Lys Phe Phe Gly Ser Lys Thr Arg Lys Ser Phe
    450                 455                 460
Phe Met Arg Gly Ser Lys Thr Ala Gln Ala Ser Ala Ser Asp Ser Glu
465                 470                 475                 480
Asp Asp Ala Ser Lys Asn Pro Gln Leu Leu Glu Gln Thr Lys Arg Leu
                485                 490                 495
Ser Gln Asn Leu Pro Val Asp Leu Phe Asp Glu His Val Asp Pro Leu
            500                 505                 510
His Arg Gln Arg Ala Leu Ser Ala Val Ser Ile Leu Thr Ile Thr Met
        515                 520                 525
Gln Glu Gln Glu Lys Phe Gln Glu Pro Cys Phe Pro Cys Gly Lys Asn
    530                 535                 540
Leu Ala Ser Lys Tyr Leu Val Trp Asp Cys Ser Pro Gln Trp Leu Cys
545                 550                 555                 560
Ile Lys Lys Val Leu Arg Thr Ile Met Thr Asp Pro Phe Thr Glu Leu
                565                 570                 575
Ala Ile Thr Ile Cys Ile Ile Ile Asn Thr Val Phe Leu Ala Val Glu
            580                 585                 590
His His Asn Met Asp Asp Asn Leu Lys Thr Ile Leu Lys Ile Gly Asn
        595                 600                 605
```

Trp Val Phe Thr Gly Ile Phe Ile Ala Glu Met Cys Leu Lys Ile Ile
610 615 620
Ala Leu Asp Pro Tyr His Tyr Phe Arg His Gly Trp Asn Val Phe Asp
625 630 635 640
Ser Ile Val Ala Leu Leu Ser Leu Ala Asp Val Leu Tyr Asn Thr Leu
645 650 655
Ser Asp Asn Asn Arg Ser Phe Leu Ala Ser Leu Arg Val Leu Arg Val
660 665 670
Phe Lys Leu Ala Lys Ser Trp Pro Thr Leu Asn Thr Leu Ile Lys Ile
675 680 685
Ile Gly His Ser Val Gly Ala Leu Gly Asn Leu Thr Val Val Leu Thr
690 695 700
Ile Val Val Phe Ile Phe Ser Val Val Gly Met Arg Leu Phe Gly Thr
705 710 715 720
Lys Phe Asn Lys Thr Ala Tyr Ala Thr Gln Glu Arg Pro Arg Arg Arg
725 730 735
Trp His Met Asp Asn Phe Tyr His Ser Phe Leu Val Val Phe Arg Ile
740 745 750
Leu Cys Gly Glu Trp Ile Glu Asn Met Trp Gly Cys Met Gln Asp Met
755 760 765
Asp Gly Ser Pro Leu Cys Ile Ile Val Phe Val Leu Ile Met Val Ile
770 775 780
Gly Lys Leu Val Val Leu Asn Leu Phe Ile Ala Leu Leu Leu Asn Ser
785 790 795 800
Phe Ser Asn Glu Glu Lys Asp Gly Ser Leu Glu Gly Glu Thr Arg Lys
805 810 815
Thr Lys Val Gln Leu Ala Leu Asp Arg Phe Arg Arg Ala Phe Ser Phe
820 825 830
Met Leu His Ala Leu Gln Ser Phe Cys Cys Lys Lys Cys Arg Arg Lys
835 840 845
Asn Ser Pro Lys Pro Lys Glu Thr Thr Glu Ser Phe Ala Gly Glu Asn
850 855 860
Lys Asp Ser Ile Leu Pro Asp Ala Arg Pro Trp Lys Glu Tyr Asp Thr
865 870 875 880
Asp Met Ala Leu Tyr Thr Gly Gln Ala Gly Ala Pro Leu Ala Pro Leu
885 890 895
Ala Glu Val Glu Asp Asp Val Glu Tyr Cys Gly Glu Gly Gly Ala Leu
900 905 910
Pro Thr Ser Gln His Ser Ala Gly Val Gln Ala Gly Asp Leu Pro Pro
915 920 925
Glu Thr Lys Gln Leu Thr Ser Pro Asp Asp Gln Gly Val Glu Met Glu
930 935 940
Val Phe Ser Glu Glu Asp Leu His Leu Ser Ile Gln Ser Pro Arg Lys
945 950 955 960
Lys Ser Asp Ala Val Ser Met Leu Ser Glu Cys Ser Thr Ile Asp Leu
965 970 975
Asn Asp Ile Phe Arg Asn Leu Gln Lys Thr Val Ser Pro Lys Lys Gln
980 985 990
Pro Asp Arg Cys Phe Pro Lys Gly Leu Ser Cys His Phe Leu Cys His
995 1000 1005
Lys Thr Asp Lys Arg Lys Ser Pro Trp Val Leu Trp Trp Asn Ile Arg
1010 1015 1020
Lys Thr Cys Tyr Gln Ile Val Lys His Ser Trp Phe Glu Ser Phe Ile
1025 1030 1035 1040

Ile Phe Val Ile Leu Leu Ser Ser Gly Ala Leu Ile Phe Glu Asp Val
1045 1050 1055
Asn Leu Pro Ser Arg Pro Gln Val Glu Lys Leu Leu Arg Cys Thr Asp
1060 1065 1070
Asn Ile Phe Thr Phe Ile Phe Leu Leu Glu Met Ile Leu Lys Trp Val
1075 1080 1085
Ala Phe Gly Phe Arg Arg Tyr Phe Thr Ser Ala Trp Cys Trp Leu Asp
1090 1095 1100
Phe Leu Ile Val Val Val Ser Val Leu Ser Leu Met Asn Leu Pro Ser
1105 1110 1115 1120
Leu Lys Ser Phe Arg Thr Leu Arg Ala Leu Arg Pro Leu Arg Ala Leu
1125 1130 1135
Ser Gln Phe Glu Gly Met Lys Val Val Val Tyr Ala Leu Ile Ser Ala
1140 1145 1150
Ile Pro Ala Ile Leu Asn Val Leu Leu Val Cys Leu Ile Phe Trp Leu
1155 1160 1165
Val Phe Cys Ile Leu Gly Val Asn Leu Phe Ser Gly Lys Phe Gly Arg
1170 1175 1180
Cys Ile Asn Gly Thr Asp Ile Asn Met Tyr Leu Asp Phe Thr Glu Val
1185 1190 1195 1200
Pro Asn Arg Ser Gln Cys Asn Ile Ser Asn Tyr Ser Trp Lys Val Pro
1205 1210 1215
Gln Val Asn Phe Asp Asn Val Gly Asn Ala Tyr Leu Ala Leu Leu Gln
1220 1225 1230
Val Ala Thr Tyr Lys Gly Trp Leu Glu Ile Met Asn Ala Ala Val Asp
1235 1240 1245
Ser Arg Glu Lys Asp Glu Gln Pro Asp Phe Glu Ala Asn Leu Tyr Ala
1250 1255 1260
Tyr Leu Tyr Phe Val Val Phe Ile Ile Phe Gly Ser Phe Phe Thr Leu
1265 1270 1275 1280
Asn Leu Phe Ile Gly Val Ile Ile Asp Asn Phe Asn Gln Gln Gln Lys
1285 1290 1295
Lys Leu Gly Gly Gln Asp Ile Phe Met Thr Glu Glu Gln Lys Lys Tyr
1300 1305 1310
Tyr Asn Ala Met Lys Lys Leu Gly Thr Lys Lys Pro Gln Lys Pro Ile
1315 1320 1325
Pro Arg Pro Leu Asn Lys Cys Gln Ala Phe Val Phe Asp Leu Val Thr
1330 1335 1340
Ser Gln Val Phe Asp Val Ile Ile Leu Gly Leu Ile Val Leu Asn Met
1345 1350 1355 1360
Ile Ile Met Met Ala Glu Ser Ala Asp Gln Pro Lys Asp Val Lys Lys
1365 1370 1375
Thr Phe Asp Ile Leu Asn Ile Ala Phe Val Val Ile Phe Thr Ile Glu
1380 1385 1390
Cys Leu Ile Lys Val Phe Ala Leu Arg Gln His Tyr Phe Thr Asn Gly
1395 1400 1405
Trp Asn Leu Phe Asp Cys Val Val Val Val Leu Ser Ile Ile Ser Thr
1410 1415 1420
Leu Val Ser Arg Leu Glu Asp Ser Asp Ile Ser Phe Pro Pro Thr Leu
1425 1430 1435 1440
Phe Arg Val Val Arg Leu Ala Arg Ile Gly Arg Ile Leu Arg Leu Val
1445 1450 1455
Arg Ala Ala Arg Gly Ile Arg Thr Leu Leu Phe Ala Leu Met Met Ser
1460 1465 1470

Leu Pro Ser Leu Phe Asn Ile Gly Leu Leu Leu Phe Leu Val Met Phe
1475 1480 1485
Ile Tyr Ala Ile Phe Gly Met Ser Trp Phe Ser Lys Val Lys Lys Gly
1490 1495 1500
Ser Gly Ile Asp Asp Ile Phe Asn Phe Glu Thr Phe Thr Gly Ser Met
1505 1510 1515 1520
Leu Cys Leu Phe Gln Ile Thr Thr Ser Ala Gly Trp Asp Thr Leu Leu
1525 1530 1535
Asn Pro Met Leu Glu Ala Lys Glu His Cys Asn Ser Ser Ser Gln Asp
1540 1545 1550
Ser Cys Gln Gln Pro Gln Ile Ala Val Val Tyr Phe Val Ser Tyr Ile
1555 1560 1565
Ile Ile Ser Phe Leu Ile Val Val Asn Met Tyr Ile Ala Val Ile Leu
1570 1575 1580
Glu Asn Phe Asn Thr Ala Thr Glu Glu Ser Glu Asp Pro Leu Gly Glu
1585 1590 1595 1600
Asp Asp Phe Glu Ile Phe Tyr Glu Val Trp Glu Lys Phe Asp Pro Glu
1605 1610 1615
Ala Ser Gln Phe Ile Gln Tyr Ser Ala Leu Ser Asp Phe Ala Asp Ala
1620 1625 1630
Leu Pro Glu Pro Leu Arg Val Ala Lys Pro Asn Lys Phe Gln Phe Leu
1635 1640 1645
Val Met Asp Leu Pro Met Val Met Gly Asp Arg Leu His Cys Met Asp
1650 1655 1660
Val Leu Phe Ala Phe Thr Thr Arg Val Leu Gly Asp Ser Ser Gly Leu
1665 1670 1675 1680
Asp Thr Met Lys Thr Met Met Glu Glu Lys Phe Met Glu Ala Asn Pro
1685 1690 1695
Phe Lys Lys Leu Tyr Glu Pro Ile Val Thr Thr Thr Lys Arg Lys Glu
1700 1705 1710
Glu Glu Gln Gly Ala Ala Val Ile Gln Arg Ala Tyr Arg Lys His Met
1715 1720 1725
Glu Lys Met Val Lys Leu Arg Leu Lys Asp Arg Ser Ser Ser Ser His
1730 1735 1740
Gln Val Phe Cys Asn Gly Asp Leu Ser Ser Leu Asp Val Ala Lys Val
1745 1750 1755 1760
Lys Val His Asn Asp
1765

(4) INFORMATION FOR SEQ ID NO:3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 856 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(F) TISSUE TYPE: Dorsal root ganglia
(G) CELL TYPE: Peripheral nerve
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCTGAGCAGT GGGGCACTGA TATTTGAAGA TGTTCACCTT GAGAACCAAC CCAAAATCCA 60

AGAATTACTA AATTGTACTG ACATTATTTT TACACATATT TTTATCCTGG AGATGGTACT 120

AAAATGGGTA GCCTTCGGAT TTGGAAAGTA TTTCACCAGT GCCTGGTGCT GCCTTGATTT 180

CATCATTGTG ATTGTCTCTG TGACCACCCT CATTAACTTA ATGGAATTGA AGTCCTTCCG 240

GACTCTACGA GCACTGAGGC CTCTTCGTGC GCTGTCCCAG TTTGAAGGAA TGAAGGTGGT 300

GGTCAATGCT CTCATAGGTG CCATACCTGC CATTCTGAAT GTTTTGCTTG TCTGCCTCAT 360

TTTCTGGCTC GTATTTTGTA TTCTGGGAGT ATACTTCTTT TCTGGAAAAT TTGGGAAATG 420

CATTAATGGA ACAGACTCAG TTATAAATTA TACCATCATT ACAAATAAAA GTCAATGTGA 480

AAGTGGCAAT TTCTCTTGGA TCAACCAGAA AGTCAACTTT GACAATGTGG GAAATGCTTA 540

CCTCGCTCTG CTGCAAGTGG CAACATTTAA GGGCTGGATG GATATTATAT ATGCAGCTGT 600

TGATTCCACA GAGAAAGAAC AACAGCCAGA GTTTGAGAGC AATTCACTCG GTTACATTTA 660

CTTCGTAGTC TTTATCATCT TTGGCTCATT CTTCACTCTG AATCTCTTCA TTGGCGTTAT 720

CATTGACAAC TTCAACCAAC AGCAGAAAAA GTTAGGTGGC CAAGACATTT TTATGACAGA 780

AGAACAGAAG AAATACTATA ATGCAATGAA AAAATTAGGA TCCAAAAAAC CTCAAAAACC 840

CATTCCACGG CCCGTT 856

(5) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 701 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RT-PCR
(A) DESCRIPTION: /desc = DNA probe/domain IV"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: rat
(F) TISSUE TYPE: dorsal root ganglia
(G) CELL TYPE: peripheral nerve

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTCAACATGG TTACGATGAT GGTGGAGACC GACGAGCAGG GCGAGGAGAA GACGAAGGTT 60

CTGGGCAGAA TCAACCAGTT CTTTGTGGCC GTCTTCACGG GCGAGTGTGT GATGAAGATG 120

TTCGCCCTGC GACAGTACTA TTTCACCAAC GGCTGGAACG TGTTCGACTT CATAGTGGTG 180

ATCCTGTCCA TTGGGAGTCT GCTGTTTCTG CAATCCTTAA GTCACTGGAA AACTACTTCT 240

CCCCGACGCT CTTCCGGGTC ATCCGTCTGG CCAGGATCGG CCGCATCCTC AGGCTGATCC 300

GAGCAGCCAA GGGGATTCGC ACGCTGCTCT TCGCCCTCAT GATGTCCCTG CCCGCCCTCT 360

TCAACATCGG CCTCCTCCTC TTCCTCGTCA TGTTCATCTA CTCCATCTTC GGCATGGCCA 420

GCTTCGCTAA CGTCGTGGAC GAGGCCGGCA TCGACGACAT GTTCAACTTC AAGACCTTTG 480

GCAACAGCAT GCTGTGCCTG TTCCAGATCA CCACCTCGGC CGGCTGGGAC GGCCTCCTCA 540

GCCCCATCCT CAACACGGGG CCTCCCTACT GCGACCCCAA CCTGCCCAAC AGCAACGGCT 600

CCCGGGGGAA CTGCGGGAGC CCGGCGGTGG GCATCATCTT CTTCACCACC TACATCATCA 660

TCTCCTTCCT CATCGTGGTC AACATGTATA TCGCAGTCAT C 701

(5) INFORMATION FOR SEQ ID NO:5:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5334 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RT-PCR
(A) DESCRIPTION: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(vi) ORIGINAL SOURCE:
(A) ORGANISM:
(F) TISSUE TYPE:
(G) CELL TYPE:
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTCGACTCTA GATCAGGGTG AAGATGGAGG AGAGGTACTA CCCGGTGATC TTCCCGGACG 60

AGCGGAATTT CCGCCCCTTC ACTTCCGACT CTCTGGCTGC CATAGAGAAG CGGATTGCTA 120

TCCAAAAGGA GAGGAAGAAG TCCAAAGACA AGGCGGCAGC TGAGCCCCAG CCTCGGCCTC 180

AGCTTGACCT AAAGGCCTCC AGGAAGTTAC CTAAGCTTTA TGGTGACATT CCCCCTGAGC 240

TTGTAGCGAA GCCTCTGGAA GACCTGGACC CATTCTACAA AGACCATAAG ACATTCATGG 300

TGTTGAACAA GAAGAGAACA ATTTATCGCT TCAGCGCCAA GCGGGCCTTG TTCATTCTGG 360

GGCCTTTTAA TCCCCTCAGA AGCTTAATGA TTCGTATCTC TGTCCATTCA GTCTTTAGCA 420

TGTTCATCAT CTGCACGGTG ATCATCAACT GTATGTTCAT GGCGAATTCT ATGGAGAGAA 480

GTTTCGACAA CGACATTCCC GAATACGTCT TCATTGGGAT TTATATTTTA GAAGCTGTGA 540

TTAAAATATT GGCAAGAGGC TTCATTGTGG ATGAGTTTTC CTTCCTCCGA GATCCGTGGA 600

ACTGGCTGGA CTTCATTGTC ATTGGAACAG CGATCGCAAC TTGTTTTCCG GGCAGCCAAG 660

TCAATCTTTC AGCTCTTCGT ACCTTCCGAG TGTTCAGAGC TCTGAAGGCG ATTTCAGTTA 720

TCTCAGGTCT GAAGGTCATC GTAGGTGCCC TGCTGCGCTC GGTGAAGAAG CTGGTAGACG 780

TGATGGTCCT CACTCTCTTC TGCCTCAGCA TCTTTGCCCT GGTCGGTCAG CAGCTGTTCA 840

TGGGAATTCT GAACCAGAAG TGTATTAAGC ACAACTGTGG CCCCAACCCT GCATCCAACA 900

AGGATTGCTT TGAAAAGGAA AAAGATAGCG AAGACTTCAT AATGTGTGGT ACCTGGCTCG 960

GCAGCAGACC CTGTCCCAAT GGTTCTACGT GCGATAAAAC CACATTGAAC CCAGACAATA 1020

ATTATACAAA GTTTGACAAC TTTGGCTGGT CCTTTCTCGC CATGTTCCGG GTTATGACTC 1080

AAGACTCCTG GGAGAGGCTT TACCGACAGA TCCTGCGGAC CTCTGGGATC TACTTTGTCT 1140

TCTTCTTCGT GGTGGTCATC TTCCTGGGCT CCTTCTACCT GCTTAACCTA ACCCTGGCTG 1200

TTGTCACCAT GGCTTATGAA GAACAGAACA GAAATGTAGC TGCTGAGACA GAGGCCAAGG 1260

AGAAAATGTT TCAGGAAGCC CAGCAGCTGT TAAGGGAGGA GAAGGAGGCT CTGGTTGCCA 1320

TGGGAATTGA CAGAAGTTCC CTTAATTCCC TTCAAGCTTC ATCCTTTTCC CCGAAGAAGA 1380

GGAAGTTTTT CGGTAGTAAG ACAAGAAAGT CCTTCTTTAT GAGAGGGTCC AAGACGGCCC 1440

AAGCCTCAGC GTCTGATTCA GAGGACGATG CCTCTAAAAA TCCACAGCTC CTTGAGCAGA 1500

CCAAACGACT GTCCCAGAAC TTGCCAGTGG ATCTCTTTGA TGAGCACGTG GACCCCCTCC 1560

ACAGGCAGAG AGCGCTGAGC GCTGTCAGTA TCTTAACCAT CACCATGCAG GAACAAGAAA 1620

AATTCCAGGA GCCTTGTTTC CCATGTGGGA AAAATTTGGC CTCTAAGTAC CTGGTGTGGG 1680

ACTGTAGCCC TCAGTGGCTG TGCATAAAGA AGGTCCTGCG GACCATCATG ACGGATCCCT 1740

TTACTGAGCT GGCCATCACC ATCTGCATCA TCATCAATAC CGTTTTCTTA GCCGTGGAGC 1800

ACCACAACAT GGATGACAAC TTAAAGACCA TACTGAAAAT AGGAAACTGG GTTTTCACGG 1860

GAATTTTCAT AGCGGAAATG TGTCTCAAGA TCATCGCGCT CGACCCTTAC CACTACTTCC 1920

GGCACGGCTG GAATGTTTTT GACAGCATCG TGGCCCTCCT GAGTCTCGCT GATGTGCTCT 1980

ACAACACACT GTCTGATAAC AATAGGTCTT TCTTGGCTTC CCTCAGAGTG CTGAGGGTCT 2040

TCAAGTTAGC CAAATCCTGG CCCACGTTAA ACACTCTCAT TAAGATCATC GGCCACTCCG 2100

TGGGCGCGCT TGGAAACCTG ACTGTGGTCC TGACTATCGT GGTCTTCATC TTTTCTGTGG 2160

TGGGCATGCG GCTCTTCGGC ACCAAGTTTA ACAAGACCGC CTACGCCACC CAGGAGCGGC 2220

CCAGGCGGCG CTGGCACATG GATAATTTCT ACCACTCCTT CCTGGTGGTG TTCCGCATCC 2280

TCTGTGGGGA ATGGATCGAG AACATGTGGG GCTGCATGCA GGATATGGAC GGCTCCCCGT 2340

TGTGCATCAT TGTCTTTGTC CTGATAATGG TGATCGGGAA GCTTGTGGTG CTTAACCTCT 2400

TCATTGCCTT GCTGCTCAAT TCCTTCAGCA ATGAGGAGAA GGATGGGAGC CTGGAAGGAG 2460

AGACCAGGAA AACCAAAGTG CAGCTAGCCC TGGATCGGTT CCGCCGGGCC TTCTCCTTCA 2520

TGCTGCACGC TCTTCAGAGT TTTTGTTGCA AGAAATGCAG GAGGAAAAAC TCGCCAAAGC 2580

CAAAAGAGAC AACAGAAAGC TTTGCTGGTG AGAATAAAGA CTCAATCCTC CCGGATGCGA 2640

GGCCCTGGAA GGAGTATGAT ACAGACATGG CTTTGTACAC TGGACAGGCC GGGGCTCCGC 2700

TGGCCCCACT CGCAGAGGTA GAGGACGATG TGGAATATTG TGGTGAAGGC GGTGCCCTAC 2760

CCACCTCACA ACATAGTGCT GGAGTTCAGG CCGGTGACCT CCCTCCAGAG ACCAAGCAGC 2820

TCACTAGCCC GGATGACCAA GGGGTTGAAA TGGAAGTATT TTCTGAAGAA GATCTGCATT 2880

TAAGCATACA GAGTCCTCGA AAGAAGTCTG ACGCAGTGAG CATGCTCTCG GAATGCAGCA 2940

CAATTGACCT GAATGATATC TTTAGAAATT TACAGAAAAC AGTTTCCCCC AAAAAGCAGC 3000

CAGATAGATG CTTTCCCAAG GGCCTTAGTT GTCACTTTCT ATGCCACAAA ACAGACAAGA 3060

GAAAGTCCCC CTGGGTCCTG TGGTGGAACA TTCGGAAAAC CTGCTACCAA ATCGTGAAGC 3120

ACAGCTGGTT TGAGAGTTTC ATAATCTTTG TTATTCTGCT GAGCAGTGGA GCGCTGATAT 3180

TTGAAGATGT CAATCTCCCC AGCCGGCCCC AAGTTGAGAA ATTACTAAGG TGTACCGATA 3240

ATATTTTCAC ATTTATTTTC CTCCTGGAAA TGATCCTGAA GTGGGTGGCC TTTGGATTCC 3300

GGAGGTATTT CACCAGTGCC TGGTGCTGGC TTGATTTCCT CATTGTGGTG GTGTCTGTGC 3360

TCAGTCTCAT GAATCTACCA AGCTTGAAGT CCTTCCGGAC TCTGCGGGCC CTGAGACCTC 3420

TGCGGGCGCT GTCCCAGTTT GAAGGAATGA AGGTTGTCGT CTACGCCCTG ATCAGCGCCA 3480

TACCTGCCAT TCTCAATGTC TTGCTGGTCT GCCTCATTTT CTGGCTCGTA TTTTGTATCT 3540

TGGGAGTAAA TTTATTTTCT GGGAAGTTTG GAAGGTGCAT TAACGGGACA GACATAAATA 3600

TGTATTTGGA TTTTACCGAA GTTCCGAACC GAAGCCAATG TAACATTAGT AATTACTCGT 3660

GGAAGGTCCC GCAGGTCAAC TTTGACAACG TGGGGAATGC CTATCTCGCC CTGCTGCAAG 3720

TGGCAACCTA TAAGGGCTGG CTGGAAATCA TGAATGCTGC TGTCGATTCC AGAGAGAAAG 3780

ACGAGCAGCC GGACTTTGAG GCGAACCTCT ACGCGTATCT CTACTTTGTG GTTTTTATCA 3840

TCTTCGGCTC CTTCTTTACC CTGAACCTCT TTATCGGTGT TATTATTGAC AACTTCAATC 3900

AGCAGCAGAA AAAGTTAGGT GGCCAAGACA TCTTCATGAC TGAGGAGCAG AAGAAATATT 3960

ACAATGCAAT GAAAAAGTTA GGAACCAAGA AACCTCAAAA GCCCATCCCA AGGCCCCTGA 4020

ACAAATGTCA AGCCTTTGTG TTCGACCTGG TCACAAGCCA GGTCTTTGAC GTCATCATTC 4080

TGGGTCTTAT TGTCTTAAAT ATGATTATCA TGATGGCTGA ATCTGCCGAC CAGCCCAAAG 4140

ATGTGAAGAA AACCTTTGAT ATCCTCAACA TAGCCTTCGT GGTCATCTTT ACCATAGAGT 4200

GTCTCATCAA AGTCTTTGCT TTGAGGCAAC ACTACTTCAC CAATGGCTGG AACTTATTTG 4260

ATTGTGTGGT CGTGGTTCTT TCTATCATTA GTACCCTGGT TTCCCGCTTG GAGGACAGTG 4320

ACATTTCTTT CCCGCCCACG CTCTTCAGAG TCGTCCGCTT GGCTCGGATT GGTCGAATCC 4380

TCAGGCTGGT CCGGGCTGCC CGGGGAATCA GGACCCTCCT CTTTGCTTTG ATGATGTCTC 4440

TCCCCTCTCT CTTCAACATC GGTCTGCTGC TCTTCCTGGT GATGTTCATT TACGCCATCT 4500

TTGGGATGAG CTGGTTTTCC AAAGTGAAGA AGGGCTCCGG GATCGACGAC ATCTTCAACT 4560

TCGAGACCTT TACGGGCAGC ATGCTGTGCC TCTTCCAGAT AACCACTTCG GCTGGCTGGG 4620

ATACCCTCCT CAACCCCATG CTGGAGGCAA AAGAACACTG CAACTCCTCC TCCCAAGACA 4680

GCTGTCAGCA GCCGCAGATA GCCGTCGTCT ACTTCGTCAG TTACATCATC ATCTCCTTCC 4740

TCATCGTGGT CAACATGTAC ATCGCTGTGA TCCTCGAGAA CTTCAACACA GCCACGGAGG 4800

AGAGCGAGGA CCCTCTGGGA GAGGACGACT TTGAAATCTT CTATGAGGTC TGGGAGAAGT 4860

TTGACCCCGA GGCGTCGCAG TTCATCCAGT ATTCGGCCCT CTCTGACTTT GCGGACGCCC 4920

TGCCGGAGCC GTTGCGTGTG GCCAAGCCGA ATAAGTTTCA GTTTCTAGTG ATGGACTTGC 4980

CCATGGTGAT GGGCGACCGC CTCCATTGCA TGGATGTTCT CTTTGCTTTC ACTACCAGGG 5040

TCCTCGGGGA CTCCAGCGGC TTGGATACCA TGAAAACCAT GATGGAGGAG AAGTTTATGG 5100

AGGCCAACCC TTTTAAGAAG CTCTACGAGC CCATAGTCAC CACCACCAAG AGGAAGGAGG 5160

AGGAGCAAGG CGCCGCCGTC ATCCAGAGGG CCTACCGGAA ACACATGGAG AAGATGGTCA 5220

AACTGAGGCT GAAGGACAGG TCAAGTTCAT CGCACCAGGT GTTTTGCAAT GGAGACTTGT 5280

CCAGCTTGGA TGTGGCCAAG GTCAAGGTTC ACAATGACTG AACCCTCATC TAGA 5334

CLAIMS

What is claimed is:

1. An isolated DNA sequence comprising the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:3.
2. The DNA of Claim 1 wherein said DNA sequence is encoding a sodium channel protein or fragment thereof.
3. The DNA of Claim 2 wherein said sodium channel protein is the α-subunit or fragment thereof.
4. The DNA of Claim 3 wherein said sodium channel protein is tetrodotoxin-resistant.
5. The DNA of Claim 3 or 4 wherein said sodium channel protein is found in mammals.
6. The DNA of Claim 3 or 4 wherein said sodium channel protein is found in rat.
7. The DNA of Claim 3 or 4 wherein said sodium channel protein is found in human.
8. The DNA of Claim 1 wherein said DNA is cDNA.
9. The DNA of Claim 1 wherein said DNA is synthetic DNA.
10. Expression vectors comprising the DNA of Claim 8.
11. Expression vectors comprising the synthetic DNA of Claim 9.
12. Host cells transformed with the expression vectors of Claim 10.
13. Host cells transformed with the expression vectors of Claim 11.
14. A recombinant polynucleotide comprising a nucleic acid sequence derived from the DNA sequence of Claim 1.
15. A sodium channel protein encoded by a DNA of Claims 1 to 9 or allelic variants thereof.
16. A tetrodotoxin-resistant sodium channel protein encoded by a DNA of Claims 1 to 9 or allelic variants thereof.
17. The protein of Claim 16 having the amino acid sequence set forth in SEQ ID NO:2.
18. A method for identifying inhibitors of tetrodotoxin-resistant sodium channel protein comprising contacting a compound suspected of being said inhibitor with sodium channel protein of claim 16 and measuring the activity of said expressed sodium channel protein.
19. Poly- and/or monoclonal antibodies raised against a tetrodotoxin-resistant sodium channel protein encoded by a DNA of Claims 1 to 9 or allelic variants thereof.
20. A diagnostic kit comprising a polynucleotide of claim 14 capable of specifically hybridizing to a tetrodotoxin-resistant sodium channel protein or fragment thereof.
21. The use of an isolated DNA sequence of Claims 1 to 9 for identifying a compound suspected of being an inhibitor of tetrodotoxin-resistant sodium channel protein.
22. The invention substantially as hereinbefore described especially with reference to the foregoing Examples.

Figure 1A: SEQ ID NO:1

```
   1  GAAGTCACAG GAGTGTCTGT CAGCGAGAGG AAGAAGGGAG AGTTTACTGA
  51  GTGTCTTCTG CCCCTCCTCA GGGTGAAGAT GGAGGAGAGG TACTACCCGG
 101  TGATCTTCCC GGACGAGCGG AATTTCCGCC CCTTCACTTC CGACTCTCTG
 151  GCTGCCATAG AGAAGCGGAT TGCTATCCAA AAGGAGAGGA AGAAGTCCAA
 201  AGACAAGGCG GCAGCTGAGC CCCAGCCTCG GCCTCAGCTT GACCTAAAGG
 251  CCTCCAGGAA GTTACCTAAG CTTTATGGTG ACATTCCCCC TGAGCTTGTA
 301  GCGAAGCCTC TGGAAGACCT GGACCCATTC TACAAAGACC ATAAGACATT
 351  CATGGTGTTG AACAAGAAGA GAACAATTTA TCGCTTCAGC GCCAAGCGGG
 401  CCTTGTTCAT TCTGGGGCCT TTTAATCCCC TCAGAAGCTT AATGATTCGT
 451  ATCTCTGTCC ATTCAGTCTT TAGCATGTTC ATCATCTGCA CGGTGATCAT
 501  CAACTGTATG TTCATGGCGA ATTCTATGGA GAGAAGTTTC GACAACGACA
 551  TTCCCGAATA CGTCTTCATT GGGATTTATA TTTTAGAAGC TGTGATTAAA
 601  ATATTGGCAA GAGGCTTCAT TGTGGATGAG TTTTCCTTCC TCCGAGATCC
 651  GTGGAACTGG CTGGACTTCA TTGTCATTGG AACAGCGATC GCAACTTGTT
 701  TTCCGGGCAG CCAAGTCAAT CTTTCAGCTC TTCGTACCTT CCGAGTGTTC
 751  AGAGCTCTGA AGGCGATTTC AGTTATCTCA GGTCTGAAGG TCATCGTAGG
 801  TGCCCTGCTG CGCTCGGTGA AGAAGCTGGT AGACGTGATG GTCCTCACTC
 851  TCTTCTGCCT CAGCATCTTT GCCCTGGTCG GTCAGCAGCT GTTCATGGGA
 901  ATTCTGAACC AGAAGTGTAT TAAGCACAAC TGTGGCCCCA ACCCTGCATC
 951  CAACAAGGAT TGCTTTGAAA AGGAAAAAGA TAGCGAAGAC TTCATAATGT
1001  GTGGTACCTG GCTCGGCAGC AGACCCTGTC CCAATGGTTC TACGTGCGAT
1051  AAAACCACAT TGAACCCAGA CAATAATTAT ACAAAGTTTG ACAACTTTGG
1101  CTGGTCCTTT CTCGCCATGT TCCGGGTTAT GACTCAAGAC TCCTGGGAGA
1151  GGCTTTACCG ACAGATCCTG CGGACCTCTG GGATCTACTT TGTCTTCTTC
1201  TTCGTGGTGG TCATCTTCCT GGGCTCCTTC TACCTGCTTA ACCTAACCCT
```

Figure 1B: SEQ ID NO:1

```
1251  GGCTGTTGTC ACCATGGCTT ATGAAGAACA GAACAGAAAT GTAGCTGCTG
1301  AGACAGAGGC CAAGGAGAAA ATGTTTCAGG AAGCCCAGCA GCTGTTAAGG
1351  GAGGAGAAGG AGGCTCTGGT TGCCATGGGA ATTGACAGAA GTTCCCTTAA
1401  TTCCCTTCAA GCTTCATCCT TTTCCCCGAA GAAGAGGAAG TTTTTCGGTA
1451  GTAAGACAAG AAAGTCCTTC TTTATGAGAG GGTCCAAGAC GGCCCAAGCC
1501  TCAGCGTCTG ATTCAGAGGA CGATGCCTCT AAAAATCCAC AGCTCCTTGA
1551  GCAGACCAAA CGACTGTCCC AGAACTTGCC AGTGGATCTC TTTGATGAGC
1601  ACGTGGACCC CCTCCACAGG CAGAGAGCGC TGAGCGCTGT CAGTATCTTA
1651  ACCATCACCA TGCAGGAACA AGAAAAATTC CAGGAGCCTT GTTTCCCATG
1701  TGGGAAAAAT TTGGCCTCTA AGTACCTGGT GTGGGACTGT AGCCCTCAGT
1751  GGCTGTGCAT AAAGAAGGTC CTGCGGACCA TCATGACGGA TCCCTTTACT
1801  GAGCTGGCCA TCACCATCTG CATCATCATC AATACCGTTT TCTTAGCCGT
1851  GGAGCACCAC AACATGGATG ACAACTTAAA GACCATACTG AAAATAGGAA
1901  ACTGGGTTTT CACGGGAATT TTCATAGCGG AAATGTGTCT CAAGATCATC
1951  GCGCTCGACC CTTACCACTA CTTCCGGCAC GGCTGGAATG TTTTTGACAG
2001  CATCGTGGCC CTCCTGAGTC TCGCTGATGT GCTCTACAAC ACACTGTCTG
2051  ATAACAATAG GTCTTTCTTG GCTTCCCTCA GAGTGCTGAG GGTCTTCAAG
2101  TTAGCCAAAT CCTGGCCCAC GTTAAACACT CTCATTAAGA TCATCGGCCA
2151  CTCCGTGGGC GCGCTTGGAA ACCTGACTGT GGTCCTGACT ATCGTGGTCT
2201  TCATCTTTTC TGTGGTGGGC ATGCGGCTCT TCGGCACCAA GTTTAACAAG
2251  ACCGCCTACG CCACCCAGGA GCGGCCCAGG CGGCGCTGGC ACATGGATAA
2301  TTTCTACCAC TCCTTCCTGG TGGTGTTCCG CATCCTCTGT GGGGAATGGA
2351  TCGAGAACAT GTGGGGCTGC ATGCAGGATA TGGACGGCTC CCCGTTGTGC
2401  ATCATTGTCT TTGTCCTGAT AATGGTGATC GGGAAGCTTG TGGTGCTTAA
```

Figure 1C: SEQ ID NO:1

2451 CCTCTTCATT GCCTTGCTGC TCAATTCCTT CAGCAATGAG GAGAAGGATG

2501 GGAGCCTGGA AGGAGAGACC AGGAAAACCA AAGTGCAGCT AGCCCTGGAT

2551 CGGTTCCGCC GGGCCTTCTC CTTCATGCTG CACGCTCTTC AGAGTTTTTG

2601 TTGCAAGAAA TGCAGGAGGA AAAACTCGCC AAAGCCAAAA GAGACAACAG

2651 AAAGCTTTGC TGGTGAGAAT AAAGACTCAA TCCTCCCGGA TGCGAGGCCC

2701 TGGAAGGAGT ATGATACAGA CATGGCTTTG TACACTGGAC AGGCCGGGGC

2751 TCCGCTGGCC CCACTCGCAG AGGTAGAGGA CGATGTGGAA TATTGTGGTG

2801 AAGGCGGTGC CCTACCCACC TCACAACATA GTGCTGGAGT TCAGGCCGGT

2851 GACCTCCCTC CAGAGACCAA GCAGCTCACT AGCCCGGATG ACCAAGGGGT

2901 TGAAATGGAA GTATTTTCTG AAGAAGATCT GCATTTAAGC ATACAGAGTC

2951 CTCGAAAGAA GTCTGACGCA GTGAGCATGC TCTCGGAATG CAGCACAATT

3001 GACCTGAATG ATATCTTTAG AAATTTACAG AAAACAGTTT CCCCCAAAAA

3051 GCAGCCAGAT AGATGCTTTC CCAAGGGCCT TAGTTGTCAC TTTCTATGCC

3101 ACAAAACAGA CAAGAGAAAG TCCCCCTGGG TCCTGTGGTG GAACATTCGG

3151 AAAACCTGCT ACCAAATCGT GAAGCACAGC TGGTTTGAGA GTTTCATAAT

3201 CTTTGTTATT CTGCTGAGCA GTGGAGCGCT GATATTTGAA GATGTCAATC

3251 TCCCCAGCCG GCCCCAAGTT GAGAAATTAC TAAGGTGTAC CGATAATATT

3301 TTCACATTTA TTTTCCTCCT GGAAATGATC CTGAAGTGGG TGGCCTTTGG

3351 ATTCCGGAGG TATTTCACCA GTGCCTGGTG CTGGCTTGAT TTCCTCATTG

3401 TGGTGGTGTC TGTGCTCAGT CTCATGAATC TACCAAGCTT GAAGTCCTTC

3451 CGGACTCTGC GGGCCCTGAG ACCTCTGCGG GCGCTGTCCC AGTTTGAAGG

3501 AATGAAGGTT GTCGTCTACG CCCTGATCAG CGCCATACCT GCCATTCTCA

3551 ATGTCTTGCT GGTCTGCCTC ATTTTCTGGC TCGTATTTTG TATCTTGGGA

3601 GTAAATTTAT TTTCTGGGAA GTTTGGAAGG TGCATTAACG GGACAGACAT

Caroline Smith Et Henderson

Figure 1D: SEQ ID NO:1

```
3651  AAATATGTAT TTGGATTTTA CCGAAGTTCC GAACCGAAGC CAATGTAACA
3701  TTAGTAATTA CTCGTGGAAG GTCCCGCAGG TCAACTTTGA CAACGTGGGG
3751  AATGCCTATC TCGCCCTGCT GCAAGTGGCA ACCTATAAGG GCTGGCTGGA
3801  AATCATGAAT GCTGCTGTCG ATTCCAGAGA GAAAGACGAG CAGCCGGACT
3851  TTGAGGCGAA CCTCTACGCG TATCTCTACT TTGTGGTTTT TATCATCTTC
3901  GGCTCCTTCT TTACCCTGAA CCTCTTTATC GGTGTTATTA TTGACAACTT
3951  CAATCAGCAG CAGAAAAAGT TAGGTGGCCA AGACATTTTT ATGACAGAAG
4001  AACAGAAGAA ATATTACAAT GCAATGAAAA AGTTAGGAAC CAAGAAACCT
4051  CAAAAGCCCA TCCCAAGGCC CCTGAACAAA TGTCAAGCCT TTGTGTTCGA
4101  CCTGGTCACA AGCCAGGTCT TTGACGTCAT CATTCTGGGT CTTATTGTCT
4151  TAAATATGAT TATCATGATG GCTGAATCTG CCGACCAGCC CAAAGATGTG
4201  AAGAAAACCT TTGATATCCT CAACATAGCC TTCGTGGTCA TCTTTACCAT
4251  AGAGTGTCTC ATCAAAGTCT TTGCTTTGAG GCAACACTAC TTCACCAATG
4301  GCTGGAACTT ATTTGATTGT GTGGTCGTGG TTCTTTCTAT CATTAGTACC
4351  CTGGTTTCCC GCTTGGAGGA CAGTGACATT TCTTTCCCGC CCACGCTCTT
4401  CAGAGTCGTC CGCTTGGCTC GGATTGGTCG AATCCTCAGG CTGGTCCGGG
4451  CTGCCCGGGG AATCAGGACC CTCCTCTTTG CTTTGATGAT GTCTCTCCCC
4501  TCTCTCTTCA ACATCGGTCT GCTGCTCTTC CTGGTGATGT TCATTTACGC
4551  CATCTTTGGG ATGAGCTGGT TTTCCAAAGT GAAGAAGGGC TCCGGGATCG
4601  ACGACATCTT CAACTTCGAG ACCTTTACGG GCAGCATGCT GTGCCTCTTC
4651  CAGATAACCA CTTCGGCTGG CTGGGATACC CTCCTCAACC CCATGCTGGA
4701  GGCAAAAGAA CACTGCAACT CCTCCTCCCA AGACAGCTGT CAGCAGCCGC
4751  AGATAGCCGT CGTCTACTTC GTCAGTTACA TCATCATCTC CTTCCTCATC
4801  GTGGTCAACA TGTACATCGC TGTGATCCTC GAGAACTTCA ACACAGCCAC
```

Figure 1E: SEQ ID NO: 1

4851 GGAGGAGAGC GAGGACCCTC TGGGAGAGGA CGACTTTGAA ATCTTCTATG

4901 AGGTCTGGGA GAAGTTTGAC CCCGAGGCGT CGCAGTTCAT CCAGTATTCG

4951 GCCCTCTCTG ACTTTGCGGA CGCCCTGCCG GAGCCGTTGC GTGTGGCCAA

5001 GCCGAATAAG TTTCAGTTTC TAGTGATGGA CTTGCCCATG GTGATGGGCG

5051 ACCGCCTCCA TTGCATGGAT GTTCTCTTTG CTTTCACTAC CAGGGTCCTC

5101 GGGGACTCCA GCGGCTTGGA TACCATGAAA ACCATGATGG AGGAGAAGTT

5151 TATGGAGGCC AACCCTTTTA AGAAGCTCTA CGAGCCCATA GTCACCACCA

5201 CCAAGAGGAA GGAGGAGGAG CAAGGCGCCG CCGTCATCCA GAGGGCCTAC

5251 CGGAAACACA TGGAGAAGAT GGTCAAACTG AGGCTGAAGG ACAGGTCAAG

5301 TTCATCGCAC CAGGTGTTTT GCAATGGAGA CTTGTCCAGC TTGGATGTGG

5351 CCAAGGTCAA GGTTCACAAT GAC**TGA**ACCC TCATCTCCAC CCCTACCTCA

5401 CTGCCTCACA GCTTAGCCTC CAGCCTCTGG CGAGCAGGCG GCAGACTCAC

5451 TGAACACAGG CCGTTCGATC TGTGTTTTTG GCTGAACGAG GTGACAGGTT

5501 GGCGTCCATT TTTAAATGAC TCTTGGAAAG ATTTCATGTA GAGAGATGTT

5551 AGAAGGGACT GCAAAGGACA CCGACCATAA CGGAAGGCCT GGAGGACAGT

5601 CCAACTTACA TAAAGATGAG AAACAAGAAG GAAAGATCCC AGGAAAACTT

5651 CAGATTGTGT TCTCAGTACA TCCCCCAATG TGTCTGTTCG GTGTTTTGAG

5701 TATGTGACCT GCCACATGTA GCTCTTTTTT GCATGTACGT CAAAACCCTG

5751 CAGTAAGTTG ATAGCTTGCT ACGGGTGTTC CTACCAGCAT CACAGAATTG

5801 GGTGTATGAC TCAAACCTAA AAGCATGACT CTGACTTGTC AGTCAGCACC

5851 CCGACTTTCA GACGCTCCAA TCTCTGTCCC AGGTGTCTAA CGAATAAATA

5901 GGTAAAAG

Figure 2A: SEQ ID NO: 2

```
Met Glu Glu Arg Tyr Tyr Pro Val Ile Phe Pro Asp Glu Arg Asn Phe
1               5                  10                  15
Arg Pro Phe Thr Ser Asp Ser Leu Ala Ala Ile Glu Lys Arg Ile Ala
            20                  25                  30
Ile Gln Lys Glu Arg Lys Lys Ser Lys Asp Lys Ala Ala Ala Glu Pro
        35                  40                  45
Gln Pro Arg Pro Gln Leu Asp Leu Lys Ala Ser Arg Lys Leu Pro Lys
    50                  55                  60
Leu Tyr Gly Asp Ile Pro Pro Glu Leu Val Ala Lys Pro Leu Glu Asp
65                  70                  75                  80
Leu Asp Pro Phe Tyr Lys Asp His Lys Thr Phe Met Val Leu Asn Lys
                85                  90                  95
Lys Arg Thr Ile Tyr Arg Phe Ser Ala Lys Arg Ala Leu Phe Ile Leu
            100                 105                 110
Gly Pro Phe Asn Pro Leu Arg Ser Leu Met Ile Arg Ile Ser Val His
        115                 120                 125
Ser Val Phe Ser Met Phe Ile Ile Cys Thr Val Ile Ile Asn Cys Met
    130                 135                 140
Phe Met Ala Asn Ser Met Glu Arg Ser Phe Asp Asn Asp Ile Pro Glu
145                 150                 155                 160
Tyr Val Phe Ile Gly Ile Tyr Ile Leu Glu Ala Val Ile Lys Ile Leu
                165                 170                 175
Ala Arg Gly Phe Ile Val Asp Glu Phe Ser Phe Leu Arg Asp Pro Trp
            180                 185                 190
Asn Trp Leu Asp Phe Ile Val Ile Gly Thr Ala Ile Ala Thr Cys Phe
        195                 200                 205
Pro Gly Ser Gln Val Asn Leu Ser Ala Leu Arg Thr Phe Arg Val Phe
    210                 215                 220
Arg Ala Leu Lys Ala Ile Ser Val Ile Ser Gly Leu Lys Val Ile Val
225                 230                 235                 240
Gly Ala Leu Leu Arg Ser Val Lys Lys Leu Val Asp Val Met Val Leu
                245                 250                 255
Thr Leu Phe Cys Leu Ser Ile Phe Ala Leu Val Gly Gln Gln Leu Phe
            260                 265                 270
Met Gly Ile Leu Asn Gln Lys Cys Ile Lys His Asn Cys Gly Pro Asn
        275                 280                 285
```

Pro Ala Ser Asn Lys Asp Cys Phe Glu Lys Glu Lys Asp Ser Glu Asp
290 295 300
Phe Ile Met Cys Gly Thr Trp Leu Gly Ser Arg Pro Cys Pro Asn Gly
305 310 315 320

Figure 2B: SEQ ID NO: 2

Ser Thr Cys Asp Lys Thr Thr Leu Asn Pro Asp Asn Asn Tyr Thr Lys
325 330 335

Phe Asp Asn Phe Gly Trp Ser Phe Leu Ala Met Phe Arg Val Met Thr
340 345 350

Gln Asp Ser Trp Glu Arg Leu Tyr Arg Gln Ile Leu Arg Thr Ser Gly
355 360 365

Ile Tyr Phe Val Phe Phe Phe Val Val Val Ile Phe Leu Gly Ser Phe
370 375 380

Tyr Leu Leu Asn Leu Thr Leu Ala Val Val Thr Met Ala Tyr Glu Glu
385 390 395 400

Gln Asn Arg Asn Val Ala Ala Glu Thr Glu Ala Lys Glu Lys Met Phe
405 410 415

Gln Glu Ala Gln Gln Leu Leu Arg Glu Glu Lys Glu Ala Leu Val Ala
420 425 430

Met Gly Ile Asp Arg Ser Ser Leu Asn Ser Leu Gln Ala Ser Ser Phe
435 440 445

Ser Pro Lys Lys Arg Lys Phe Phe Gly Ser Lys Thr Arg Lys Ser Phe
450 455 460

Phe Met Arg Gly Ser Lys Thr Ala Gln Ala Ser Ala Ser Asp Ser Glu
465 470 475 480

Asp Asp Ala Ser Lys Asn Pro Gln Leu Leu Glu Gln Thr Lys Arg Leu
485 490 495

Ser Gln Asn Leu Pro Val Asp Leu Phe Asp Glu His Val Asp Pro Leu
500 505 510

His Arg Gln Arg Ala Leu Ser Ala Val Ser Ile Leu Thr Ile Thr Met
515 520 525

Gln Glu Gln Glu Lys Phe Gln Glu Pro Cys Phe Pro Cys Gly Lys Asn
530 535 540

Leu Ala Ser Lys Tyr Leu Val Trp Asp Cys Ser Pro Gln Trp Leu Cys
545 550 555 560

Ile Lys Lys Val Leu Arg Thr Ile Met Thr Asp Pro Phe Thr Glu Leu
565 570 575

Ala Ile Thr Ile Cys Ile Ile Ile Asn Thr Val Phe Leu Ala Val Glu
580 585 590

His His Asn Met Asp Asp Asn Leu Lys Thr Ile Leu Lys Ile Gly Asn

595 600 605

Trp Val Phe Thr Gly Ile Phe Ile Ala Glu Met Cys Leu Lys Ile Ile

610 615 620

Figure 2C: SEQ ID NO: 2

```
Ala Leu Asp Pro Tyr His Tyr Phe Arg His Gly Trp Asn Val Phe Asp
625                 630                 635                 640
Ser Ile Val Ala Leu Leu Ser Leu Ala Asp Val Leu Tyr Asn Thr Leu
                645                 650                 655
Ser Asp Asn Asn Arg Ser Phe Leu Ala Ser Leu Arg Val Leu Arg Val
            660                 665                 670
Phe Lys Leu Ala Lys Ser Trp Pro Thr Leu Asn Thr Leu Ile Lys Ile
        675                 680                 685
Ile Gly His Ser Val Gly Ala Leu Gly Asn Leu Thr Val Val Leu Thr
    690                 695                 700
Ile Val Val Phe Ile Phe Ser Val Val Gly Met Arg Leu Phe Gly Thr
705                 710                 715                 720
Lys Phe Asn Lys Thr Ala Tyr Ala Thr Gln Glu Arg Pro Arg Arg Arg
                725                 730                 735
Trp His Met Asp Asn Phe Tyr His Ser Phe Leu Val Val Phe Arg Ile
            740                 745                 750
Leu Cys Gly Glu Trp Ile Glu Asn Met Trp Gly Cys Met Gln Asp Met
        755                 760                 765
Asp Gly Ser Pro Leu Cys Ile Ile Val Phe Val Leu Ile Met Val Ile
    770                 775                 780
Gly Lys Leu Val Val Leu Asn Leu Phe Ile Ala Leu Leu Leu Asn Ser
785                 790                 795                 800
Phe Ser Asn Glu Glu Lys Asp Gly Ser Leu Glu Gly Glu Thr Arg Lys
                805                 810                 815
Thr Lys Val Gln Leu Ala Leu Asp Arg Phe Arg Arg Ala Phe Ser Phe
            820                 825                 830
Met Leu His Ala Leu Gln Ser Phe Cys Cys Lys Lys Cys Arg Arg Lys
        835                 840                 845
Asn Ser Pro Lys Pro Lys Glu Thr Thr Glu Ser Phe Ala Gly Glu Asn
    850                 855                 860
Lys Asp Ser Ile Leu Pro Asp Ala Arg Pro Trp Lys Glu Tyr Asp Thr
865                 870                 875                 880
Asp Met Ala Leu Tyr Thr Gly Gln Ala Gly Ala Pro Leu Ala Pro Leu
                885                 890                 895
```

Ala Glu Val Glu Asp Asp Val Glu Tyr Cys Gly Glu Gly Gly Ala Leu
900 905 910

Figure 2D: SEQ ID NO: 2

Pro Thr Ser Gln His Ser Ala Gly Val Gln Ala Gly Asp Leu Pro Pro
915 920 925

Glu Thr Lys Gln Leu Thr Ser Pro Asp Asp Gln Gly Val Glu Met Glu
930 935 940

Val Phe Ser Glu Glu Asp Leu His Leu Ser Ile Gln Ser Pro Arg Lys
945 950 955 960

Lys Ser Asp Ala Val Ser Met Leu Ser Glu Cys Ser Thr Ile Asp Leu
965 970 975

Asn Asp Ile Phe Arg Asn Leu Gln Lys Thr Val Ser Pro Lys Lys Gln
980 985 990

Pro Asp Arg Cys Phe Pro Lys Gly Leu Ser Cys His Phe Leu Cys His
995 1000 1005

Lys Thr Asp Lys Arg Lys Ser Pro Trp Val Leu Trp Trp Asn Ile Arg
1010 1015 1020

Lys Thr Cys Tyr Gln Ile Val Lys His Ser Trp Phe Glu Ser Phe Ile
1025 1030 1035 1040

Ile Phe Val Ile Leu Leu Ser Ser Gly Ala Leu Ile Phe Glu Asp Val
1045 1050 1055

Asn Leu Pro Ser Arg Pro Gln Val Glu Lys Leu Leu Arg Cys Thr Asp
1060 1065 1070

Asn Ile Phe Thr Phe Ile Phe Leu Leu Glu Met Ile Leu Lys Trp Val
1075 1080 1085

Ala Phe Gly Phe Arg Arg Tyr Phe Thr Ser Ala Trp Cys Trp Leu Asp
1090 1095 1100

Phe Leu Ile Val Val Val Ser Val Leu Ser Leu Met Asn Leu Pro Ser
1105 1110 1115 1120

Leu Lys Ser Phe Arg Thr Leu Arg Ala Leu Arg Pro Leu Arg Ala Leu
1125 1130 1135

Ser Gln Phe Glu Gly Met Lys Val Val Val Tyr Ala Leu Ile Ser Ala
1140 1145 1150

Ile Pro Ala Ile Leu Asn Val Leu Leu Val Cys Leu Ile Phe Trp Leu
1155 1160 1165

Val Phe Cys Ile Leu Gly Val Asn Leu Phe Ser Gly Lys Phe Gly Arg
1170 1175 1180

Cys Ile Asn Gly Thr Asp Ile Asn Met Tyr Leu Asp Phe Thr Glu Val

1185 1190 1195 1200

Pro Asn Arg Ser Gln Cys Asn Ile Ser Asn Tyr Ser Trp Lys Val Pro

1205 1210 1215

Figure 2E: SEQ ID NO: 2

Gln Val Asn Phe Asp Asn Val Gly Asn Ala Tyr Leu Ala Leu Leu Gln
1220 1225 1230

Val Ala Thr Tyr Lys Gly Trp Leu Glu Ile Met Asn Ala Ala Val Asp
1235 1240 1245

Ser Arg Glu Lys Asp Glu Gln Pro Asp Phe Glu Ala Asn Leu Tyr Ala
1250 1255 1260

Tyr Leu Tyr Phe Val Val Phe Ile Ile Phe Gly Ser Phe Phe Thr Leu
1265 1270 1275 1280

Asn Leu Phe Ile Gly Val Ile Ile Asp Asn Phe Asn Gln Gln Gln Lys
1285 1290 1295

Lys Leu Gly Gly Gln Asp Ile Phe Met Thr Glu Glu Gln Lys Lys Tyr
1300 1305 1310

Tyr Asn Ala Met Lys Lys Leu Gly Thr Lys Lys Pro Gln Lys Pro Ile
1315 1320 1325

Pro Arg Pro Leu Asn Lys Cys Gln Ala Phe Val Phe Asp Leu Val Thr
1330 1335 1340

Ser Gln Val Phe Asp Val Ile Ile Leu Gly Leu Ile Val Leu Asn Met
1345 1350 1355 1360

Ile Ile Met Met Ala Glu Ser Ala Asp Gln Pro Lys Asp Val Lys Lys
1365 1370 1375

Thr Phe Asp Ile Leu Asn Ile Ala Phe Val Val Ile Phe Thr Ile Glu
1380 1385 1390

Cys Leu Ile Lys Val Phe Ala Leu Arg Gln His Tyr Phe Thr Asn Gly
1395 1400 1405

Trp Asn Leu Phe Asp Cys Val Val Val Val Leu Ser Ile Ile Ser Thr
1410 1415 1420

Leu Val Ser Arg Leu Glu Asp Ser Asp Ile Ser Phe Pro Pro Thr Leu
1425 1430 1435 1440

Phe Arg Val Val Arg Leu Ala Arg Ile Gly Arg Ile Leu Arg Leu Val
1445 1450 1455

Arg Ala Ala Arg Gly Ile Arg Thr Leu Leu Phe Ala Leu Met Met Ser
1460 1465 1470

Leu Pro Ser Leu Phe Asn Ile Gly Leu Leu Leu Phe Leu Val Met Phe
1475 1480 1485

Ile Tyr Ala Ile Phe Gly Met Ser Trp Phe Ser Lys Val Lys Lys Gly

1490 1495 1500

Ser Gly Ile Asp Asp Ile Phe Asn Phe Glu Thr Phe Thr Gly Ser Met

1505 1510 1515 1520

Figure 2F: SEQ ID NO: 2

Leu Cys Leu Phe Gln Ile Thr Thr Ser Ala Gly Trp Asp Thr Leu Leu
1525 1530 1535

Asn Pro Met Leu Glu Ala Lys Glu His Cys Asn Ser Ser Ser Gln Asp
1540 1545 1550

Ser Cys Gln Gln Pro Gln Ile Ala Val Val Tyr Phe Val Ser Tyr Ile
1555 1560 1565

Ile Ile Ser Phe Leu Ile Val Val Asn Met Tyr Ile Ala Val Ile Leu
1570 1575 1580

Glu Asn Phe Asn Thr Ala Thr Glu Glu Ser Glu Asp Pro Leu Gly Glu
1585 1590 1595 1600

Asp Asp Phe Glu Ile Phe Tyr Glu Val Trp Glu Lys Phe Asp Pro Glu
1605 1610 1615

Ala Ser Gln Phe Ile Gln Tyr Ser Ala Leu Ser Asp Phe Ala Asp Ala
1620 1625 1630

Leu Pro Glu Pro Leu Arg Val Ala Lys Pro Asn Lys Phe Gln Phe Leu
1635 1640 1645

Val Met Asp Leu Pro Met Val Met Gly Asp Arg Leu His Cys Met Asp
1650 1655 1660

Val Leu Phe Ala Phe Thr Thr Arg Val Leu Gly Asp Ser Ser Gly Leu
1665 1670 1675 1680

Asp Thr Met Lys Thr Met Met Glu Glu Lys Phe Met Glu Ala Asn Pro
1685 1690 1695

Phe Lys Lys Leu Tyr Glu Pro Ile Val Thr Thr Thr Lys Arg Lys Glu
1700 1705 1710

Glu Glu Gln Gly Ala Ala Val Ile Gln Arg Ala Tyr Arg Lys His Met
1715 1720 1725

Glu Lys Met Val Lys Leu Arg Leu Lys Asp Arg Ser Ser Ser Ser His
1730 1735 1740

Gln Val Phe Cys Asn Gly Asp Leu Ser Ser Leu Asp Val Ala Lys Val
1745 1750 1755 1760

Lys Val His Asn Asp
1765

Figure 2G: SEQ ID NO:2

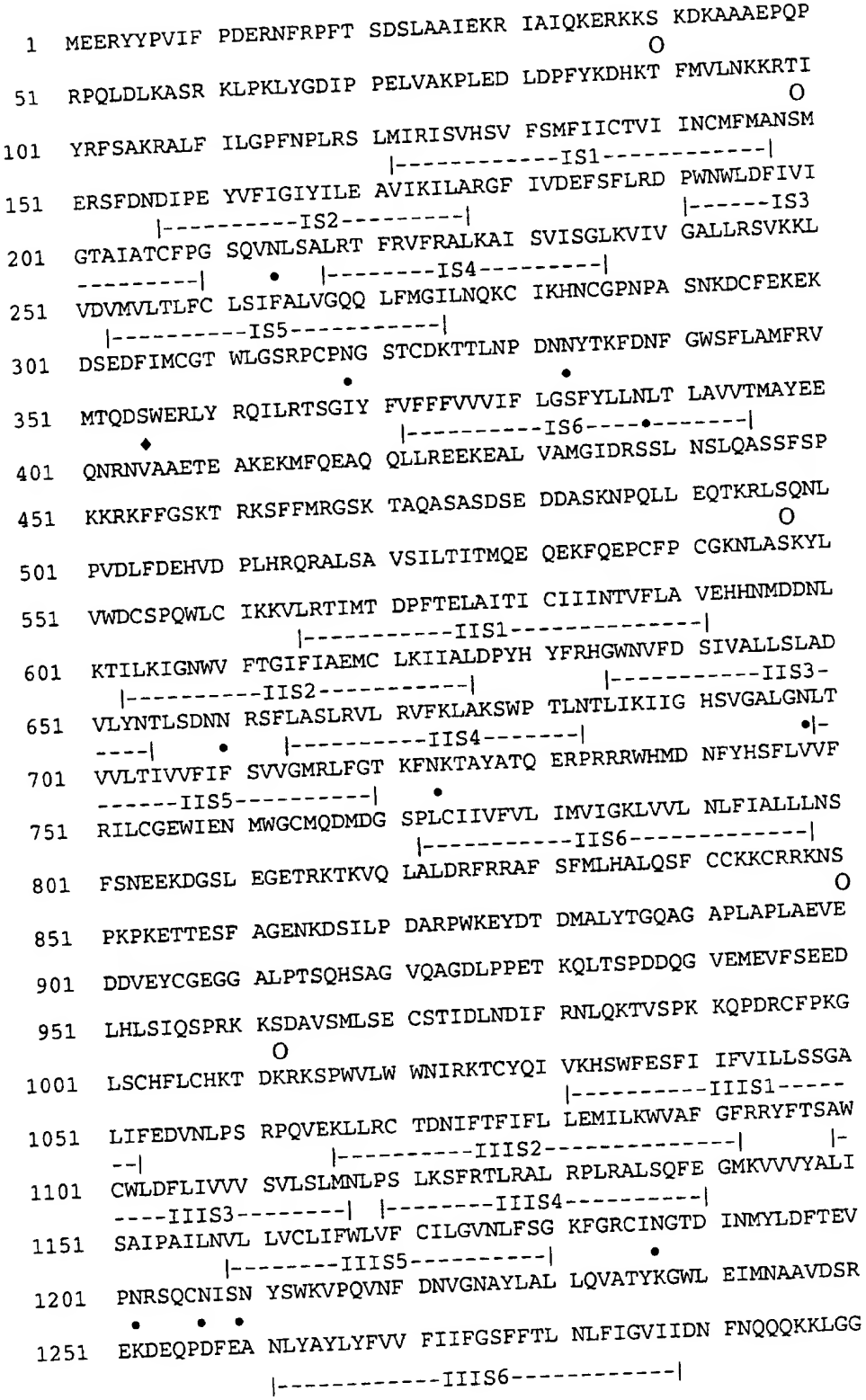


Figure 2H: SEQ ID NO: 2

```
1301  QDIFMTEEQK KYYNAMKKLG TKKPQKPIPR PLNKCQAFVF DLVTSQVFDV
                                                  |--------
1351  IILGLIVLNM IIMMAESADQ PKDVKKTFDI LNIAFVVIFT IECLIKVFAL
      IVS1------------|          |---------IVS2------------|
1401  RQHYFTNGWN LFDCVVVVLS IISTLVSRLE DSDISFPPTL FRVVRLARIG
            |----------IVS3------------|          |--------
1451  RILRLVRAAR GIRTLLFALM MSLPSLFNIG LLLFLVMFIY AIFGMSWFSK
      IVS4----------|                  |----------IVS5---------
1501  VKKGSGIDDI FNFETFTGSM LCLFQITTSA GWDTLLNPML EAKEHCNSSS
       |   O                                              •
1551  QDSCQQPQIA VVYFVSYIII SFLIVVNMYI AVILENFNTA TEESEDPLGE
             |-----------IVS6-------------|
1601  DDFEIFYEVW EKFDPEASQF IQYSALSDFA DALPEPLRVA KPNKFQFLVM

1651  DLPMVMGDRL HCMDVLFAFT TRVLGDSSGL DTMKTMMEEK FMEANPFKKL

1701  YEPIVTTTKR KEEEQGAAVI QRAYRKHMEK MVKLSLKDRS SSSHQVFCNG

1751  DLSSLDVAKV KVHND*
```

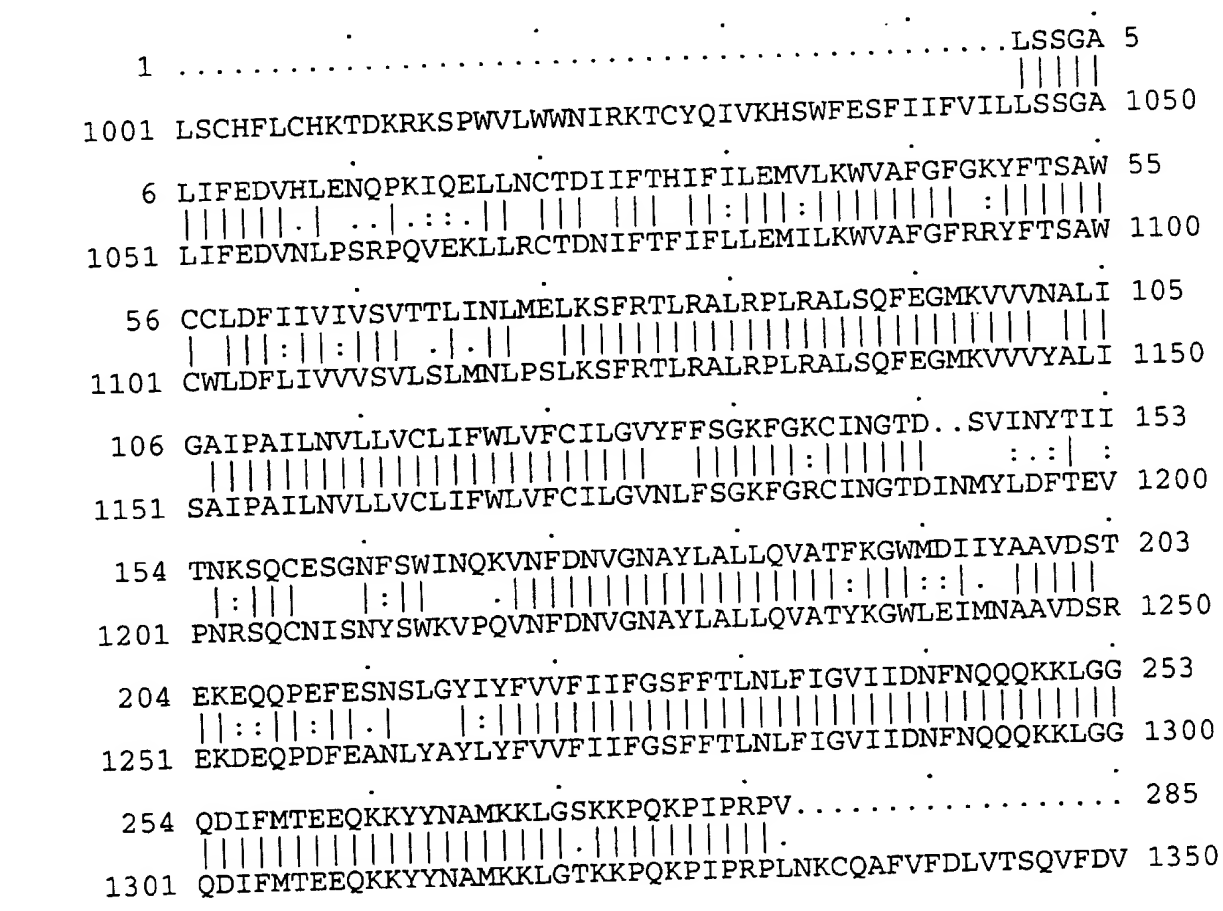
Figure 3A: SEQ ID NO:3

```
   1  GCTGAGCAGT GGGGCACTGA TATTTGAAGA TGTTCACCTT GAGAACCAAC
  51  CCAAAATCCA AGAATTACTA AATTGTACTG ACATTATTTT TACACATATT
 101  TTTATCCTGG AGATGGTACT AAAATGGGTA GCCTTCGGAT TTGGAAAGTA
 151  TTTCACCAGT GCCTGGTGCT GCCTTGATTT CATCATTGTG ATTGTCTCTG
 201  TGACCACCCT CATTAACTTA ATGGAATTGA AGTCCTTCCG GACTCTACGA
 251  GCACTGAGGC CTCTTCGTGC GCTGTCCCAG TTTGAAGGAA TGAAGGTGGT
 301  GGTCAATGCT CTCATAGGTG CCATACCTGC CATTCTGAAT GTTTTGCTTG
 351  TCTGCCTCAT TTTCTGGCTC GTATTTTGTA TTCTGGGAGT ATACTTCTTT
 401  TCTGGAAAAT TTGGGAAATG CATTAATGGA ACAGACTCAG TTATAAATTA
 451  TACCATCATT ACAAATAAAA GTCAATGTGA AAGTGGCAAT TTCTCTTGGA
 501  TCAACCAGAA AGTCAACTTT GACAATGTGG GAAATGCTTA CCTCGCTCTG
 551  CTGCAAGTGG CAACATTTAA GGGCTGGATG GATATTATAT ATGCAGCTGT
 601  TGATTCCACA GAGAAAGAAC AACAGCCAGA GTTTGAGAGC AATTCACTCG
 651  GTTACATTTA CTTCGTAGTC TTTATCATCT TTGGCTCATT CTTCACTCTG
 701  AATCTCTTCA TTGGCGTTAT CATTGACAAC TTCAACCAAC AGCAGAAAAA
 751  GTTAGGTGGC CAAGACATTT TTATGACAGA AGAACAGAAG AAATACTATA
 801  ATGCAATGAA AAAATTAGGA TCCAAAAAAC CTCAAAAACC CATTCCACGG

 851  CCCGTT
```

Caulfield Stratton Et Henderson

Figure 3B: SEQ ID NO:3



Cowling, Strathy & Henderson

Figure 4: SEQ ID NO:4

```
1     CTCAACATGG TTACGATGAT GGTGGAGACC  GACGAGCAGG  GCGAGGAGAA
51    GACGAAGGTT CTGGGCAGAA TCAACCAGTT  CTTTGTGGCC  GTCTTCACGG
101   GCGAGTGTGT GATGAAGATG TTCGCCCTGC  GACAGTACTA  TTTCACCAAC
151   GGCTGGAACG TGTTCGAcTT CATAGTGGTG  ATCCTGTCCA  TTGGGAGTCT
201   GCTGTTTCT  GCAATCCTTA AGTCACTGGA  AAACTACTTC  TCCCCGACGC
251   TCTTCCGGGT CATCCGTCTG GCCAGGATCG  GCCGCATCCT  CAGGCTGATC
301   CGAGCAGCCA AGGGGATTCG CACGCTGCTC  TTCGCCCTCA  TGATGTCCCT
351   GCCCGCCCTC TTCAACATCG GCCTCCTCCT  CTTCCTCGtC  ATGTTCATCT
401   ACTCCATCTT CGGCATGGCC AGCTTCGCTA  ACGTCGTGGA  CGAGGCCGGC
451   ATCGACGACA TGTTCAACTT CAAGACCTTT  GGCAACAGCA  TGCTGTGCCT
501   GTTCCAGATC ACCACCTCGG CCGGCTGGGA  CGGCCTCCTC  AGCCCCATCC
551   TCAACACGGG GCCTCCCTAC TGCGACCCCA  ACCTGCCCAA  CAGCAACGGC
601   TCCCGGGGGA ACTGCGGGAG CCCGGCGGTG  GGCATCATCT  TCTTCACCAC
651   CTACATCATC ATCTCCTTCC TCATCGTGGT  CAACATGTAT  ATCGCAGTCA
701   TC
```

Figure 5A: SEQ ID NO: 5

```
   1  GTCGACTCTA GATCAGGGTG AAGATGGAGG AGAGGTACTA CCCGGTGATC
  51  TTCCCGGACG AGCGGAATTT CCGCCCCTTC ACTTCCGACT CTCTGGCTGC
 101  CATAGAGAAG CGGATTGCTA TCCAAAAGGA GAGGAAGAAG TCCAAAGACA
 151  AGGCGGCAGC TGAGCCCCAG CCTCGGCCTC AGCTTGACCT AAAGGCCTCC
 201  AGGAAGTTAC CTAAGCTTTA TGGTGACATT CCCCCTGAGC TTGTAGCGAA
 251  GCCTCTGGAA GACCTGGACC CATTCTACAA AGACCATAAG ACATTCATGG
 301  TGTTGAACAA GAAGAGAACA ATTTATCGCT TCAGCGCCAA GCGGGCCTTG
 351  TTCATTCTGG GGCCTTTTAA TCCCCTCAGA AGCTTAATGA TTCGTATCTC
 401  TGTCCATTCA GTCTTTAGCA TGTTCATCAT CTGCACGGTG ATCATCAACT
 451  GTATGTTCAT GGCGAATTCT ATGGAGAGAA GTTTCGACAA CGACATTCCC
 501  GAATACGTCT TCATTGGGAT TTATATTTTA GAAGCTGTGA TTAAAATATT
 551  GGCAAGAGGC TTCATTGTGG ATGAGTTTTC CTTCCTCCGA GATCCGTGGA
 601  ACTGGCTGGA CTTCATTGTC ATTGGAACAG CGATCGCAAC TTGTTTTCCG
 651  GGCAGCCAAG TCAATCTTTC AGCTCTTCGT ACCTTCCGAG TGTTCAGAGC
 701  TCTGAAGGCG ATTTCAGTTA TCTCAGGTCT GAAGGTCATC GTAGGTGCCC
 751  TGCTGCGCTC GGTGAAGAAG CTGGTAGACG TGATGGTCCT CACTCTCTTC
 801  TGCCTCAGCA TCTTTGCCCT GGTCGGTCAG CAGCTGTTCA TGGGAATTCT
 851  GAACCAGAAG TGTATTAAGC ACAACTGTGG CCCCAACCCT GCATCCAACA
 901  AGGATTGCTT TGAAAAGGAA AAAGATAGCG AAGACTTCAT AATGTGTGGT
 951  ACCTGGCTCG GCAGCAGACC CTGTCCCAAT GGTTCTACGT GCGATAAAAC
1001  CACATTGAAC CCAGACAATA ATTATACAAA GTTTGACAAC TTTGGCTGGT
1051  CCTTTCTCGC CATGTTCCGG GTTATGACTC AAGACTCCTG GGAGAGGCTT
1101  TACCGACAGA TCCTGCGGAC CTCTGGGATC TACTTTGTCT TCTTCTTCGT
```

Figure 5B: SEQ ID NO: 5

1151 GGTGGTCATC TTCCTGGGCT CCTTCTACCT GCTTAACCTA ACCCTGGCTG

1201 TTGTCACCAT GGCTTATGAA GAACAGAACA GAAATGTAGC TGCTGAGACA

1251 GAGGCCAAGG AGAAAATGTT TCAGGAAGCC CAGCAGCTGT TAAGGGAGGA

1301 GAAGGAGGCT CTGGTTGCCA TGGGAATTGA CAGAAGTTCC CTTAATTCCC

1351 TTCAAGCTTC ATCCTTTTCC CCGAAGAAGA GGAAGTTTTT CGGTAGTAAG

1401 ACAAGAAAGT CCTTCTTTAT GAGAGGGTCC AAGACGGCCC AAGCCTCAGC

1451 GTCTGATTCA GAGGACGATG CCTCTAAAAA TCCACAGCTC CTTGAGCAGA

1501 CCAAACGACT GTCCCAGAAC TTGCCAGTGG ATCTCTTTGA TGAGCACGTG

1551 GACCCCCTCC ACAGGCAGAG AGCGCTGAGC GCTGTCAGTA TCTTAACCAT

1601 CACCATGCAG GAACAAGAAA AATTCCAGGA GCCTTGTTTC CCATGTGGGA

1651 AAAATTTGGC CTCTAAGTAC CTGGTGTGGG ACTGTAGCCC TCAGTGGCTG

1701 TGCATAAAGA AGGTCCTGCG GACCATCATG ACGGATCCCT TTACTGAGCT

1751 GGCCATCACC ATCTGCATCA TCATCAATAC CGTTTTCTTA GCCGTGGAGC

1801 ACCACAACAT GGATGACAAC TTAAAGACCA TACTGAAAAT AGGAAACTGG

1851 GTTTTCACGG GAATTTTCAT AGCGGAAATG TGTCTCAAGA TCATCGCGCT

1901 CGACCCTTAC CACTACTTCC GGCACGGCTG GAATGTTTTT GACAGCATCG

1951 TGGCCCTCCT GAGTCTCGCT GATGTGCTCT ACAACACACT GTCTGATAAC

2001 AATAGGTCTT TCTTGGCTTC CCTCAGAGTG CTGAGGGTCT TCAAGTTAGC

2051 CAAATCCTGG CCCACGTTAA ACACTCTCAT TAAGATCATC GGCCACTCCG

2101 TGGGCGCGCT TGGAAACCTG ACTGTGGTCC TGACTATCGT GGTCTTCATC

2151 TTTTCTGTGG TGGGCATGCG GCTCTTCGGC ACCAAGTTTA ACAAGACCGC

2201 CTACGCCACC CAGGAGCGGC CCAGGCGGCG CTGGCACATG GATAATTTCT

2251 ACCACTCCTT CCTGGTGGTG TTCCGCATCC TCTGTGGGGA ATGGATCGAG

2301 AACATGTGGG GCTGCATGCA GGATATGGAC GGCTCCCCGT TGTGCATCAT

Figure 5C: SEQ ID NO: 5

2351 TGTCTTTGTC CTGATAATGG TGATCGGGAA GCTTGTGGTG CTTAACCTCT

2401 TCATTGCCTT GCTGCTCAAT TCCTTCAGCA ATGAGGAGAA GGATGGGAGC

2451 CTGGAAGGAG AGACCAGGAA AACCAAAGTG CAGCTAGCCC TGGATCGGTT

2501 CCGCCGGGCC TTCTCCTTCA TGCTGCACGC TCTTCAGAGT TTTTGTTGCA

2551 AGAAATGCAG GAGGAAAAAC TCGCCAAAGC CAAAAGAGAC AACAGAAAGC

2601 TTTGCTGGTG AGAATAAAGA CTCAATCCTC CCGGATGCGA GGCCCTGGAA

2651 GGAGTATGAT ACAGACATGG CTTTGTACAC TGGACAGGCC GGGGCTCCGC

2701 TGGCCCCACT CGCAGAGGTA GAGGACGATG TGGAATATTG TGGTGAAGGC

2751 GGTGCCCTAC CCACCTCACA ACATAGTGCT GGAGTTCAGG CCGGTGACCT

2801 CCCTCCAGAG ACCAAGCAGC TCACTAGCCC GGATGACCAA GGGGTTGAAA

2851 TGGAAGTATT TTCTGAAGAA GATCTGCATT TAAGCATACA GAGTCCTCGA

2901 AAGAAGTCTG ACGCAGTGAG CATGCTCTCG GAATGCAGCA CAATTGACCT

2951 GAATGATATC TTTAGAAATT TACAGAAAAC AGTTTCCCCC AAAAAGCAGC

3001 CAGATAGATG CTTTCCCAAG GGCCTTAGTT GTCACTTTCT ATGCCACAAA

3051 ACAGACAAGA GAAAGTCCCC CTGGGTCCTG TGGTGGAACA TTCGGAAAAC

3101 CTGCTACCAA ATCGTGAAGC ACAGCTGGTT TGAGAGTTTC ATAATCTTTG

3151 TTATTCTGCT GAGCAGTGGA GCGCTGATAT TTGAAGATGT CAATCTCCCC

3201 AGCCGGCCCC AAGTTGAGAA ATTACTAAGG TGTACCGATA ATATTTTCAC

3251 ATTTATTTTC CTCCTGGAAA TGATCCTGAA GTGGGTGGCC TTTGGATTCC

3301 GGAGGTATTT CACCAGTGCC TGGTGCTGGC TTGATTTCCT CATTGTGGTG

2251 GTGTCTGTGC TCAGTCTCAT GAATCTACCA AGCTTGAAGT CCTTCCGGAC

3401 TCTGCGGGCC CTGAGACCTC TGCGGGCGCT GTCCCAGTTT GAAGGAATGA

3451 AGGTTGTCGT CTACGCCCTG ATCAGCGCCA TACCTGCCAT TCTCAATGTC

3501 TTGCTGGTCT GCCTCATTTT CTGGCTCGTA TTTTGTATCT TGGGAGTAAA

Figure 5D: SEQ ID NO: 5

3551 TTTATTTTCT GGGAAGTTTG GAAGGTGCAT TAACGGGACA GACATAAATA

3601 TGTATTTGGA TTTTACCGAA GTTCCGAACC GAAGCCAATG TAACATTAGT

3651 AATTACTCGT GGAAGGTCCC GCAGGTCAAC TTTGACAACG TGGGGAATGC

3701 CTATCTCGCC CTGCTGCAAG TGGCAACCTA TAAGGGCTGG CTGGAAATCA

3751 TGAATGCTGC TGTCGATTCC AGAGAGAAAG ACGAGCAGCC GGACTTTGAG

3801 GCGAACCTCT ACGCGTATCT CTACTTTGTG GTTTTTATCA TCTTCGGCTC

3851 CTTCTTTACC CTGAACCTCT TTATCGGTGT TATTATTGAC AACTTCAATC

3901 AGCAGCAGAA AAAGTTAGGT GGCCAAGACA TCTTCATGAC TGAGGAGCAG

3951 AAGAAATATT ACAATGCAAT GAAAAAGTTA GGAACCAAGA AACCTCAAAA

4001 GCCCATCCCA AGGCCCCTGA ACAAATGTCA AGCCTTTGTG TTCGACCTGG

4051 TCACAAGCCA GGTCTTTGAC GTCATCATTC TGGGTCTTAT TGTCTTAAAT

4101 ATGATTATCA TGATGGCTGA ATCTGCCGAC CAGCCCAAAG ATGTGAAGAA

4151 AACCTTTGAT ATCCTCAACA TAGCCTTCGT GGTCATCTTT ACCATAGAGT

4201 GTCTCATCAA AGTCTTTGCT TTGAGGCAAC ACTACTTCAC CAATGGCTGG

4251 AACTTATTTG ATTGTGTGGT CGTGGTTCTT TCTATCATTA GTACCCTGGT

4301 TTCCCGCTTG GAGGACAGTG ACATTTCTTT CCCGCCCACG CTCTTCAGAG

4351 TCGTCCGCTT GGCTCGGATT GGTCGAATCC TCAGGCTGGT CCGGGCTGCC

4401 CGGGGAATCA GGACCCTCCT CTTTGCTTTG ATGATGTCTC TCCCCTCTCT

4451 CTTCAACATC GGTCTGCTGC TCTTCCTGGT GATGTTCATT TACGCCATCT

4501 TTGGGATGAG CTGGTTTTCC AAAGTGAAGA AGGGCTCCGG GATCGACGAC

4551 ATCTTCAACT TCGAGACCTT TACGGGCAGC ATGCTGTGCC TCTTCCAGAT

4601 AACCACTTCG GCTGGCTGGG ATACCCTCCT CAACCCCATG CTGGAGGCAA

4651 AAGAACACTG CAACTCCTCC TCCCAAGACA GCTGTCAGCA GCCGCAGATA

4701 GCCGTCGTCT ACTTCGTCAG TTACATCATC ATCTCCTTCC TCATCGTGGT

Cooley Godward Et. Henderson

Figure 5E: SEQ ID NO: 5

4751 CAACATGTAC ATCGCTGTGA TCCTCGAGAA CTTCAACACA GCCACGGAGG

4801 AGAGCGAGGA CCCTCTGGGA GAGGACGACT TTGAAATCTT CTATGAGGTC

4851 TGGGAGAAGT TTGACCCCGA GGCGTCGCAG TTCATCCAGT ATTCGGCCCT

4901 CTCTGACTTT GCGGACGCCC TGCCGGAGCC GTTGCGTGTG GCCAAGCCGA

4951 ATAAGTTTCA GTTTCTAGTG ATGGACTTGC CCATGGTGAT GGGCGACCGC

5001 CTCCATTGCA TGGATGTTCT CTTTGCTTTC ACTACCAGGG TCCTCGGGGA

5051 CTCCAGCGGC TTGGATACCA TGAAAACCAT GATGGAGGAG AAGTTTATGG

5101 AGGCCAACCC TTTTAAGAAG CTCTACGAGC CCATAGTCAC CACCACCAAG

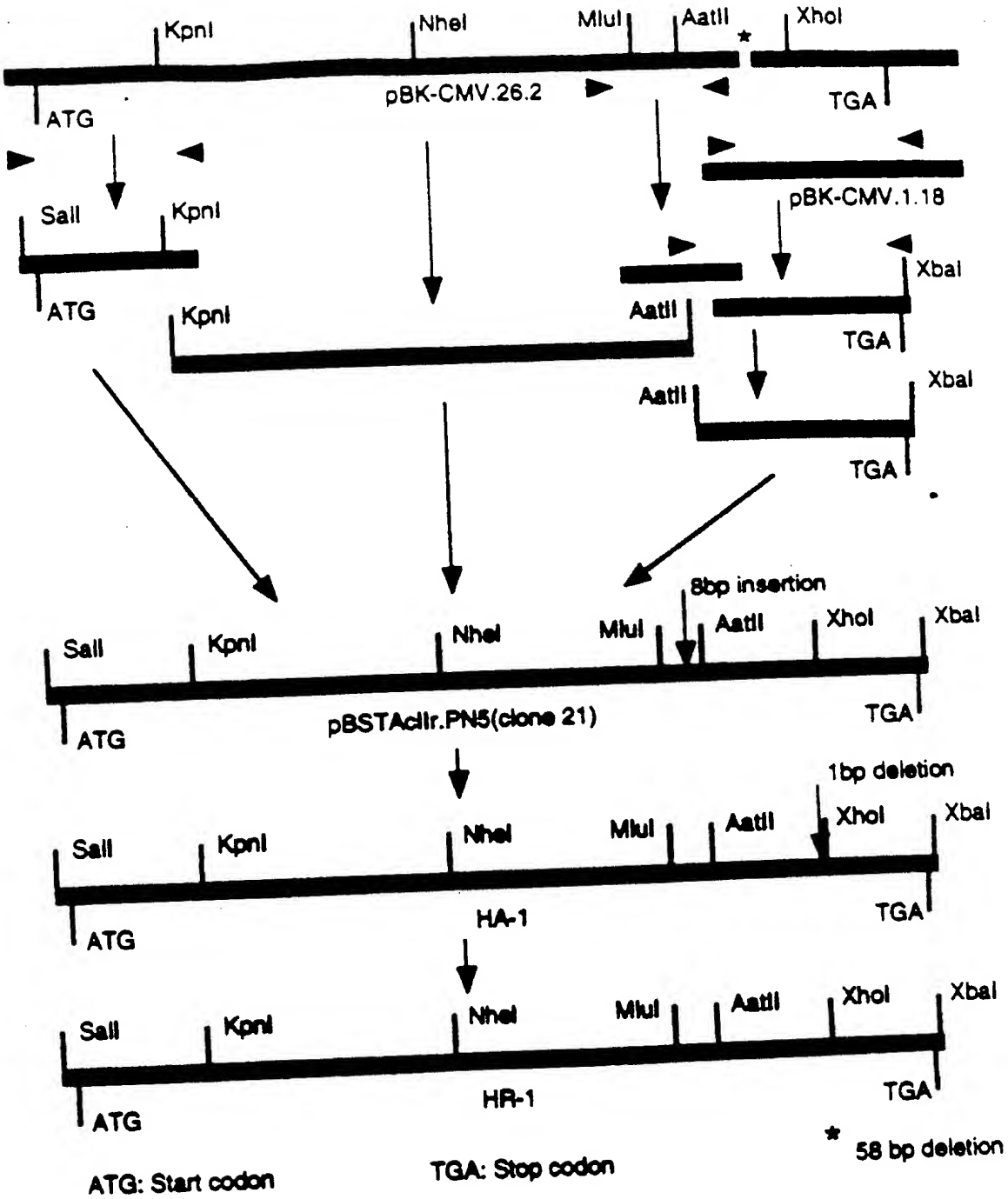
5151 AGGAAGGAGG AGGAGCAAGG CGCCGCCGTC ATCCAGAGGG CCTACCGGAA

5201 ACACATGGAG AAGATGGTCA AACTGAGGCT GAAGGACAGG TCAAGTTCAT

5251 CGCACCAGGT GTTTTGCAAT GGAGACTTGT CCAGCTTGGA TGTGGCCAAG

5301 GTCAAGGTTC ACAATGACTG AACCCTCATC TAGA

Figure 6


KpnI
NheI
MluI
AatII
XhoI
pBK-CMV.26.2
ATG
TGA
pBK-CMV.1.18
SalI
KpnI
XbaI
ATG
KpnI
AatII
TGA
AatII
XbaI
TGA
8bp insertion
SalI
KpnI
NheI
MluI
AatII
XhoI
XbaI
pBSTAclIr.PN5(clone 21)
ATG
TGA
1bp deletion
SalI
KpnI
NheI
MluI
AatII
XhoI
XbaI
HA-1
ATG
TGA
SalI
KpnI
NheI
MluI
AatII
XhoI
XbaI
HR-1
ATG
TGA
58 bp deletion
ATG: Start codon
TGA: Stop codon